

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 02:21:01 ; Search time 12121 Seconds  
(without alignments)  
11646.585 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcccgctgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb ba.\*

2: gb htg.\*

3: gb in.\*

4: gb om.\*

5: gb ov.\*

6: gb pat.\*

7: gb ph.\*

8: gb pl.\*

9: gb pr.\*

10: gb ro.\*

11: gb sts.\*

12: gb sy.\*

13: gb un.\*

14: gb vi.\*

15: em ba.\*

16: em fun.\*

17: em hum.\*

18: em in.\*

19: em mu.\*

20: em om.\*

21: em or.\*

22: em ov.\*

23: em pat.\*

24: em ph.\*

25: em pl.\*

26: em ro.\*

27: em sts.\*

28: em un.\*

29: em vi.\*

30: em htg\_hum.\*

31: em htg\_inv.\*

32: em htg\_other.\*

33: em htg\_mus.\*

34: em htg\_pln.\*

35: em htg\_rod.\*

36: em htg\_mam.\*

37: em htg\_vrt.\*

38: em sy.\*

39: em htgo\_hum.\*

40: em htgo\_mus.\*

41: em htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3030.2	93.0	3326	9	BC040679 Homo sapi
2	2696.6	82.8	190517	9	AL139022 Human chr
3	1423.8	43.7	2210	9	BSX40825 Homo sapi
4	1192.8	36.6	7924	6	AX347040 Sequence
5	1192.8	36.6	7924	6	AX348456 Sequence
6	1160.8	35.6	7924	6	AX347041 Sequence
7	1160.8	35.6	7924	6	AX348457 Sequence
8	1000.2	30.7	1054	6	AX399903 Sequence
9	866.2	26.6	895	6	AX014147 Sequence
10	866.2	26.6	895	6	BD205055 Human nuc
11	813	25.0	3139	10	BC027769 Mus muscu
12	633	19.4	676	6	AX781147 Sequence
13	577.6	17.7	945	10	RATRAB15X M83679 Sprague-Daw
14	477.4	14.7	481	6	AX396088 Sequence
15	438.8	13.5	463	6	AX334820 Sequence
16	423.6	13.0	270654	2	AC096084 Rattus no
17	416.4	12.8	2560	10	BC013790 Mus muscu
18	383	11.8	410	6	AX779778 Sequence
19	378.4	11.6	190726	2	AC132331 Mus muscu
20	338.2	10.4	2348	5	BC063736 Xenopus l
21	307	9.4	313	6	AX341683 Sequence
22	234.2	7.2	2048	9	BC002977 Homo sapi
23	234.2	7.2	2818	9	AK025165 Homo sapi
24	230.6	7.1	1980	9	S53268 Homo sapien
25	229	7.0	624	9	AF498943 Homo sapi
26	229	7.0	624	9	BT007184 Homo sapi
27	229	7.0	624	12	BT008275 Synthetic
28	229	7.0	660	9	HSRAB8 X56741 H.sapiens m
29	227.4	7.0	638	10	S53270 MEL-RAS-rel
30	224.2	6.9	1337	10	BC019990 Mus muscu
31	222.6	6.8	760	4	CFRAB8 X56385 Canine rab8
32	218.2	6.7	765	5	DYGORA2 M8391 Discopoge o
33	217.8	6.7	624	6	AX236078 Sequence
34	217.8	6.7	1128	9	BC020654 Homo sapi
35	217.8	6.7	1161	6	AX236076 Sequence
36	217.8	6.7	1265	9	AB038995 Homo sapi
37	217.8	6.7	2497	6	AX285074 Sequence
38	217.8	6.7	2497	6	AX285080 Sequence
39	217.8	6.7	2497	6	AX285089 Sequence
40	217.8	6.7	2497	6	AX285091 Sequence
41	217.8	6.7	3077	6	AX876024 Sequence
42	217.8	6.7	3077	6	BD155904 Primer fo
43	217.8	6.7	3077	9	AK001111 Homo sapi
44	216.2	6.6	3110	9	HSRAB8 AL833365 Homo sapi
45	214.2	6.6	866	6	AX866231 Sequence

ALIGNMENTS

RESULT 1  
BC040679  
LOCUS  
DEFINITION Homo sapiens RAB15, member RAS oncogene family, mRNA (CDNA clone  
MGC:42319 IMAGE:4817835), complete cds.  
ACCESSION BC040679.2 GI:34783346  
VERSION MGC.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 3326)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
PRI 11-DEC-2003  
linear mRNA 3326 bp  
BC040679  
MGC:42319 IMAGE:4817835), complete cds.

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Butterfield,Y.S., Krzywinaki,M.I., Skalska,U., Snailus,D.E.,  
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.,  
Generation and analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 3326)  
Strausberg,R.  
Direct Submission  
Submitted (29-NOV-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003 this sequence version replaced gi:26251822.  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Snailus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 70 Row: p Column: 2.  
Location/Qualifiers  
1. 3326  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="MGC:42319 IMAGE:4817835"  
/tissue\_type="Brain, hippocampus"  
/clone\_lib="NIH\_MGC\_95"  
/lab\_host="DH10B"  
/notes="vector: pbluescript"  
1. 3326  
/gene="RAB15"  
/db\_xref="LocusID:376267"  
5. -631  
/codon\_start=1  
/product="Ras-related protein Rab-15"  
/protein\_id="AAH40679.2"  
/db\_xref="GI:34783347"  
/db\_xref="LocusID:376267"  
/translaton="MAKQYDVLPRLLLLIGDSGVKTCCLCRFTDNEFHSHTIGVD  
FMKTEVDGIKVIQIWDAGERTQITIKQYRRAQGLFLYDISSERSYOHMKW  
VSDVDEVDATSLPGCGEGAPKARRGPDGKASRKLCLPQPMKTSYTHOKARR  
SLLGIRLMRSRNGRWEESKGSRRWSNMTSMKQVAPTSTLKSRRV"

misc\_feature 26.343  
/notes="RAB: Region: Rab subfamily of small GTPases"  
/db\_xref="CDD:cd00154"  
ORIGIN  
Query Match 93.0%; Score 3030.2; DB 9; Length 3326;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 3187; Conservative 0; Mismatches 8; Indels 132; Gaps 2;  
Qy 41 AGTCATGGCGAAGCAGTACGATGCTGCTCCGGCTCTGCTGATCGGGGACTCCGGGGT 100  
Db 1 AGTCATGGCGAAGCAGTACGATGCTGCTCCGGCTCTGCTGATCGGGGACTCCGGGGT 60  
Qy 101 GGCACAGACCTGCTGCTGCGCTTCACCGCAACAGAGTTCACCTCTCGCACATCTC 160  
Db 61 GGGCAAGACCTGCTGCTGCGCTTCACCGCAACAGAGTTCACCTCTCGCACATCTC 120  
Qy 161 CACCATCGGTGTTGACTTTTAAGATGAAGACCATAGAGGTAGAGGCATCAAAAGTCCGGAT 220  
Db 121 CACCATCGGTGTTGACTTTTAAGATGAAGACCATAGAGGTAGAGGCATCAAAAGTCCGGAT 180  
Qy 221 ACAGATCTGGGACACTGTCAGGGCAGGAGATACAGACCATCACAAGCAGTACTATCG 280  
Db 181 ACAGATCTGGGACACTGTCAGGGCAGGAGATACAGACCATCACAAGCAGTACTATCG 240  
Qy 281 GCGGGCCAGGGGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGACAT 340  
Db 241 GCGGGCCAGGGGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGACAT 300  
Qy 341 CATGAAGTGGGTGCTGAGTCACTGGATGA-----GTACGCCACAGAGCGCTCCAG 367  
Db 301 CATGAAGTGGGTGCTGAGTCACTGGATGAGGTAGGAGTGCACCTCACTGCCGGGTGG 360  
Qy 368-----GTACGCCACAGAGCGCTCCAG 367  
Db 361 AGAGGTGCTTACCAGGGGAAGCAAGCGAGGGCCAGATGGGAAGCAAAATGTTCCAG 420  
Qy 368-----GTACGCCACAGAGCGCTCCAG 389  
Db 421 GAAGCTTTGCTTCCAGCGCTTGAATGAAGACCTCTGTGACGCCACAGAGCGCTCCAG 480  
Qy 390 AAGATCTTATTCGGGAATAGGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGACAA 449  
Db 481 AAGATCTTATTCGGGAATAGGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGACAA 540  
Qy 450 GGCAGCAGCTGCGGAAGAGTATGCGATGGACTTCTATGAAAACAAGTGCCTGCACCAAC 509  
Db 541 GGCAGCAGCTGCGGAAGAGTATGCGATGGACTTCTATGAAAACAAGTGCCTGCACCAAC 600  
Qy 510 CTCACACATTAAAGAGTCACTTACGCGCTCTGACAGAGCTGTGTCGAGGCCCATAGGAAG 569  
Db 601 CTCACACATTAAAGAGTCACTTACGCGCTCTGACAGAGCTGTGTCGAGGCCCATAGGAAG 660  
Qy 570 GAGCTGGAAGGCTCCGGATCGGTGCGCAGCAATGAGTTGSCACTGGCAGAGCTGGAGGAG 629  
Db 661 GAGCTGGAAGGCTCCGGATCGGTGCGCAGCAATGAGTTGSCACTGGCAGAGCTGGAGGAG 720  
Qy 630 GAGGAGGCAAAACCCGAGGGCCAGGAACTCTTGAATAACCTGTGCTGCTGAGTCTTG 689  
Db 721 GAGGAGGCAAAACCCGAGGGCCAGGAACTCTTGAATAACCTGTGCTGCTGAGTCTTG 780  
Qy 690 TGTGGGGCAGCCCCACACAGCACCCCTTTCCTCAGAGGCCCTGGGAGAGAGGGGAG 749  
Db 781 TGTGGGGCAGCCCCACACAGCACCCCTTTCCTCAGAGGCCCTGGGAGAGAGGGGAG 840  
Qy 750 CCGGGGCTTTGGCCCTGCTGCTGCTCTCTGCTGATGACCCCTTATGAGTATCAGTAGCCA 809  
Db 841 CCGGGGCTTTGGCCCTGCTGCTGCTCTCTGCTGATGACCCCTTATGAGTATCAGTAGCCA 900  
Qy 810 CTACTCCCCCTGCTGGCCCTGAGAGGGCTCTGCTGCTATCTCAAGACAGCCCTGTCCC 869  
Db 901 CTACTCCCCCTGCTGGCCCTGAGAGCGGCTCTGCTGCTATCTCAAGACAGCCCTGTCCC 960

Qy	870	CAGCCGTCACCCCTGAGTGGTCTTCTTTCAGCCTGTTTCCCGACGACCAAGCGCTGCTAC	929
Db	961	CAGCCGTCACCCCTGAGTGGTCTTCTTTCAGCCTGTTTCCCGACGACCAAGCGCTGCTAC	1020
Qy	930	GACCCCAAGTGTGCCGAAGCACTGCTCTCACCATCCGACCCACCCACGACCAACAGCCA	989
Db	1021	GACCCCAAGTGTGCCGAAGCACTGCTCTCACCATCCGACCCACCCACGACCAACAGCCA	1080
Qy	990	GGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTTTCCTCCGTGTCATCGTGTCTTCTCTG	1049
Db	1081	GGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTTTCCTCCGTGTCATCGTGTCTTCTCTG	1140
Qy	1050	CTTTTTCCTCTTTCCTCCCACTTCTCTTTTCTCTGACCCCTCCCTCCGGTGGTTCGTAT	1109
Db	1141	CTTTTTCCTCTTTCCTCCCACTTCTCTTTTCTCTGACCCCTCCCTCCGGTGGTTCGTAT	1200
Qy	1110	CAAGCTCCTCAAAACCCGCTCCCGGTGTCCTGCTGTGTGTCAGCTCGCTCTTTCCTTC	1169
Db	1201	C-AAGCTCCTCAAAACCCGCTCCCGGTGTCCTGCTGTGTGTCAGCTCGCTCTTTCCTTC	1259
Qy	1170	CTTCTTAAGCTATTCGAAGGGGATGGACCCAGGCTCGTGGGAGGTTCCACCCCTTGGATCC	1229
Db	1260	CTTCTTAAGCTATTCGAAGGGGATGGACCCAGGCTCGTGGGAGGTTCCACCCCTTGGATCC	1319
Qy	1230	AGGAAGAACCCCTCCACCTGCTGGTGGGCGAAAGGCTACAGGGTGTCTTTCCTTC	1289
Db	1320	AGGAAGAACCCCTCCACCTGCTGGTGGGCGAAAGGCTACAGGGTGTCTTTCCTTC	1379
Qy	1290	TTCCCCCACCCCACTGTCTCCTCATGTGCCATGGGCTGCCTCCCACTGACCTGGCGAC	1349
Db	1380	TTCCCCCACCCCACTGTCTCCTCATGTGCCATGGGCTGCCTCCCACTGACCTGGCGAC	1439
Qy	1350	GTGGAGCATCGAGTGTAGGAGGGAACAGCAACCGGGAGTCTCTGAGCCTTGGGGCTGCC	1409
Db	1440	GTGGAGCATCGAGTGTAGGAGGGAACCGCAACCGAGGAGTCTCTGAGCCTTGGGGCTGCC	1499
Qy	1410	TACCTCTACCCATTTCCCGACAGAGCTTTTGCCCTTCTGGCTGTCGCCGCTGCCTCTTT	1469
Db	1500	TACCTCTACCCATTTCCCGACAGAGCTTTTGCCCTTCTGGCTGTCGCCGCTGCCTCTTT	1559
Qy	1470	GGGGNACTGAGCTCAGAGGCAGTGTCTCAGAGNAGGAACAAATGAGGGGTGGCAGGG	1529
Db	1560	GGGGNACTGAGCTCAGAGGCAGTGTCTCAGAGNAGGAACAAATGAGGGGTGGCAGGG	1619
Qy	1530	ATAAAAAGTCACTCCATTTCTACCTCCCATGAGCATGAACAACAATTTCTCTCCACCT	1589
Db	1620	ATAAAAAGTCACTCCATTTCTACCTCCCATGAGCATGAACAACAATTTCTCTCCACCT	1679
Qy	1590	GGCTCCCAAAATTTAAGATGTGGACCAAGGCTGTGGGTACTCCAGGGGCAAGGAGACC	1649
Db	1680	GGCTCCCAAAATTTAAGATGTGGACCAAGGCTGTGGGTACTCCAGGGGCAAGGAGACC	1739
Qy	1650	CTGGGGTCACTGACACTGTACGGCAACATGCACTCCAAAGGGGAGCATTTTGGAAAT	1709
Db	1740	CTGGGGTCACTGACACTGTGAGGCAACATGCACTCCAAAGGGGAGCATTTTGGAAAT	1799
Qy	1710	GAGGACTAGTCTCTATGTATCAGTTTAAGCAAGGAGAGCTGGCCAGGACGACAGT	1769
Db	1800	GAGGACTAGTCTCTATGTATCAGTTTAAGCAAGGAGAGCTGGCCAGGACGACAGT	1859
Qy	1770	TTGCACAGCAGAGGGGAATGTAGCAACAGAGGGCTCTTAGGGCCCATCTTCCATTCT	1829
Db	1860	TTGCACAGCAGAGGGGAATGTAGCAACAGAGGGCTCTTAGGGCCCATCTTCCATTCT	1919
Qy	1830	TAGGTAAGAGAGCATTTCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTCAGCAA	1889
Db	1920	TAGGTAAGAGAGCATTTCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTCAGCAA	1979
Qy	1890	CCAGGTTCTCTGGGACCCAAAGTTTATGGGAAGGCAAGACATTCATGGGAAGAGA	1949
Db	1980	CCAGGTTCTCTGGGACCCAAAGTTTATGGGAAGGCAAGACATTCATGGGAAGAGA	2039
Qy	1950	GAAGGAAGGGCTCGGGTAGAAACCGCTTGGTCTGCTTCTTCTTGGCTTTTAAGAACAAGCG	2009

Db	2040	GAAGGAAGGCCCTGGGTAGAAACGCTTGGTGTCTCTCTTTGGCCCTTTAAGACAAGCG	2099
Qy	2010	CTCATCTTGCCCTCTACCTCCTCATAGAGCTTGAGGGTTGCCAACACACTGTGGGTACCA	2069
Db	2100	CTCATCTTGCCCTCTACCTCCTCATAGAGCTTGAGGGTTGCCAACACACTGTGGGTACCA	2159
Qy	2070	GGTGGAGGGAAGAGACTCCCTTCTCAGAGTGTCTATGTTTCAGGAAGTTCCTTTAACCCC	2129
Db	2160	GGTGGAGGGAAGAGAGACTCCCTTCTCAGAGTGTCTATGTTTCAGGAAGTTCCTTTAACCCG	2219
Qy	2130	ATATGGCCCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGACACAGTAATTTACCAGT	2189
Db	2220	ATATGGCCCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGACACAGTAATTTACCAGT	2279
Qy	2190	TCTACTGGGTTTCTTGCCCAACCGTCCCAAGTGGGCGAGGCCCTAGGAAGAGGGTCACTTCT	2249
Db	2280	TCTACTGGGTTTCTTGCCCAACCGTCCCAAGTGGGCGAGGCCCTAGGAAGAGGGTCACTTCT	2339
Qy	2250	TAAAGCCACACATTTAGCTGTCACTCGTGGTGTGCAGGCCAAACAAAGAACTCGGTGTGAGT	2309
Db	2340	TAAAGCCACACATTTAGCTGTCACTCGTGGTGTGCAGGCCAAACAAAGAACTCGGTGTGAGT	2399
Qy	2310	ATTTCATCAACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGATAAGAACTCACTT	2369
Db	2400	ATTTCATCAACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGATAAGAACTCACTT	2459
Qy	2370	CTTTACTCTCTCAAAAGAAAGTGGGGAAGAAACCATCAAACTTCTCTCTGACTTACCA	2429
Db	2460	CTTTACTCTCTCAAAAGAAAGTGGGGAAGAAACCATCAAACTTCTCTCTGACTTACCA	2519
Qy	2430	AACAGGAAACACAGCAGGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATG	2489
Db	2520	AACAGGAAACACAGCAGGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATG	2579
Qy	2490	GTGGAAGCAAGGAGAGCAGGAAGTTGTTAAATCATCTGGCTAATGAGAAAGGAGACAGC	2549
Db	2580	GTGGAAGCAAGGAGAGCAGGAAGTTGTTAAATCATCTGGCTAATGAGAAAGGAGACAGC	2639
Qy	2550	TAACTCTAGGATCAAGCTGTGACTAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTG	2609
Db	2640	TAACTCTAGGATCAAGCTGTGACTAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTG	2699
Qy	2610	GGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGATGCCAGCCCTCA	2669
Db	2700	GGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGATGCCAGCCCTCA	2759
Qy	2670	AGTCTGTCTTCAGCAGGGACTTTGAGAACTTATTTGGGCAGTGGCTCCAACTGTGGAC	2729
Db	2760	AGTCTGTCTTCAGCAGGGACTTTGAGAACTTATTTGGGCAGTGGCTCCAACTGTGGAC	2819
Qy	2730	CAGTATTTACGCTTCCCTCAAGATCAGGCAGGGTGCCATTTCATGTCTTCTCTCCTAG	2789
Db	2820	CAGTATTTACGCTTCCCTCAAGATCAGGCAGGGTGCCATTTCATGTCTTCTCTCCTAG	2879
Qy	2790	CCCCCTCAGGAAAGAAGCACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGGAAGAAC	2849
Db	2880	CCCCCTCAGGAAAGAAGCACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGGAAGAAC	2939
Qy	2850	ATGGAATCAGGAATTCATAGACTGATAGGCCCTTATCCAAAGGCCCATGACTGGGAAAG	2909
Db	2940	ATGGAATCAGGAATTCATAGACTGATAGGCCCTTATCCAAAGGCCCATGACTGGGAAAG	2999
Qy	2910	GTATGGGACAGAAAGGAGAAATTCGGATTTTAGGGTGCACTAGCTCACCTTAACTTTT	2969
Db	3000	GTATGGGACAGAAAGGAGAAATTCGGATTTTAGGGTGCACTAGCTCACCTTAACTTTT	3059
Qy	2970	GGTGGCCTGGGGCATGTCTTGAGGCCCAGACTGTTTAAAGCAGGCTCTGCTGGCCCTGTTTAC	3029
Db	3060	GGTGGCCTGGGGCATGTCTTGAGGCCCAGACTGTTTAAAGCAGGCTCTGCTGGCCCTGTTTAC	3119
Qy	3030	TCGFTACCACTCTGCACTGTCTTGTAGACTTCATCCAGGCCCAAGGCACGCCACCTG	3089

Db	3120	TCGTCACCACTCTGCACCTGCTGTCTTGAGACTCCATCCAGCCCGGACGCCACCTG	3179	
Qy	3090	CTCCTGAGCTCCACTATCTCCCTGTGACGGGTGAACCTGCTGCTACTGTCTCGGGTCC	3149	
Db	3180	CTCCTGAGCTCCACTATCTCCCTGTGACGGGTGAACCTGCTGCTACTGTCTCGGGTCC	3239	
Qy	3150	ATATATGAATTGTGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAATAAAGAT	3209	
Db	3240	ATATATGAATTGTGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAATAAAGAT	3299	
Qy	3210	TATTTCAAACCAACCAAAAAA	3236	
Db	3300	TATTTCAAACCAACCAAAAAA	3326	
RESULT 2				
CNS01DX4		190517 bp	DNA	linear
LOCUS				PRI 17-MAY-2001
DEFINITION		Human chromosome 14 DNA sequence BAC R-840119 of library RPCI-11		
ACCESSION		from chromosome 14 of Homo sapiens (human), complete sequence.		
VERSION		AL139022		
KEYWORDS		AL139022.4	GI:14148954	
SOURCE		HTG.		
ORGANISM		Homo sapiens (human)		
REFERENCE		Homo sapiens		
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
		Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,		
		Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,		
		Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,		
		Gyapay, G., Saurin, W. and Weissenbach, J.		
		Sequencing of the human chromosome 14		
		Unpublished		
TITLE		2 (bases 1 to 190517)		
JOURNAL		Genoscope.		
AUTHORS		Direct Submission		
		Submitted (17-MAY-2001) Genoscope - Centre National de Sequencage :		
		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
		- Web : www.genoscope.cns.fr)		
		On May 18, 2001 this sequence version replaced gi:10798506.		
COMMENT		----- Genome Center		
		Center: Genoscope / Centre National de Sequencage		
		Center code: GS		
		Web site: http://www.genoscope.cns.fr/		
		Contact: SeqRef@genoscope.cns.fr		
		-----		
		The following BAC sequence is oriented from the T7 to the SP6 end.		
		Upstream BAC (overlapping the T7 end) : C-2509G16 (AC=AL355076)		
		Downstream BAC (overlapping the SP6 end) : C-2300N4 -----		
		Summary Statistics		
		Assembly program: Phrap; version 2.0		
		Quality coverage: 7.14x in Q20 bases; sum-of-contigs		
		-----		
		Overall quality chart :		
		Range : bases		
		0 : 9		
		10 - 19 : 45		
		20 - 29 : 444		
		30 - 39 : 4501		
		40 - 49 : 11521		
		50 - 59 : 12036		
		60 - 69 : 23986		
		70 - 79 : 56256		
		80 - 89 : 81728		
		90 - 99 : 81728		
		-----		
		Percentage of bases with a quality value >= 40 : 99 %.		
FEATURES		Location/Qualifiers		
source		1..190517		
		/organism="Homo sapiens"		

STS		/mol_type="genomic DNA"		
		/db_xref="taxon:9606"		
		/chromosome="14"		
		/clone="R-840119"		
		/clone_lib="RPCI-11"		
		22163..22307		
		/note="matching EMBL:H99639		
		RHdb:RH68943		
		dbSTS:STS48849		
		Identified using the e-PCR software (G. Schuler)"		
		30401..30624		
		/note="matching EMBL:R42375		
		RHdb:RH69913		
		dbSTS:STS49790		
		Identified using the e-PCR software (G. Schuler)"		
		30772..30949		
		/note="matching EMBL:AA053315		
		RHdb:RH67850		
		dbSTS:STS47777		
		Identified using the e-PCR software (G. Schuler)"		
		43552..43739		
		/note="matching EMBL:Z39082		
		RHdb:RH25206		
		dbSTS:STS25011		
		Identified using the e-PCR software (G. Schuler)"		
		44712..44852		
		/note="matching EMBL:LO0635		
		RHdb:RH53870		
		dbSTS:STS42913		
		Identified using the e-PCR software (G. Schuler)"		
		44712..44832		
		/note="matching EMBL:LO0635		
		RHdb:RH53755		
		dbSTS:STS12411		
		Identified using the e-PCR software (G. Schuler)"		
		47222..47411		
		/note="matching EMBL:AA165116		
		RHdb:RH48540		
		dbSTS:STS41599		
		Identified using the e-PCR software (G. Schuler)"		
		128322..128437		
		/note="matching EMBL:G07650		
		RHdb:RH34434		
		dbSTS:STS6376		
		Identified using the e-PCR software (G. Schuler)"		
		164135..164257		
		/note="matching EMBL:AA160869		
		RHdb:RH48489		
		dbSTS:STS41548		
		Identified using the e-PCR software (G. Schuler)"		
		169170..169323		
		/note="matching EMBL:W56308		
		RHdb:RH47847		
		dbSTS:STS40906		
		Identified using the e-PCR software (G. Schuler)"		
		171252..171384		
		/note="matching EMBL:Z39818		
		RHdb:RH10964		
		dbSTS:STS1017		
		Identified using the e-PCR software (G. Schuler)"		
		173564..173713		
		/note="matching EMBL:R10900		
		RHdb:RH53901		
		dbSTS:STS22552		
		Identified using the e-PCR software (G. Schuler)"		
		173883..174012		
		/note="matching EMBL:G27797		
		RHdb:RH32719		
		dbSTS:STS4675		
		Identified using the e-PCR software (G. Schuler)"		
ORIGIN				
Query Match		82.8%; Score 2696.6; DB 9; Length 190517;		











Db 721 TTTAAGATAAAGCGTTTATTTTGTGTTTTTATTTTGTATAGTTTGTAGGGTTGTGTAATT 780  
Qy 2056 ACACGTGGCTACAGGTGGAGGAGGAGGACTCCTCTCCAGAGTCTATGTTTCAGGAA 2115  
Db 781 ATATTGGTTATAGGTGGAGGAGGAGGATTTTTTTTTTTAGAGTGTATGTTTAGGAA 840  
Qy 2116 GTTTCCTTTAAACCCCATATGGCCCAAGAGTACCTCGTAGAGGCCCTTTAAAGACGGAACA 2175  
Db 841 GTTTCCTTTAAATTTATATGTTTAAAGAGTAGTTCGTAGGAGGTTTAAAGACGGAATA 900  
Qy 2176 AGTAATTTTACAGTCTACTGGGTTCTCTGCCACCGTCCCAAGTGGCGAGCGCTAGG 2235  
Db 901 AGTAATTTATTTAGTTTATTTGGGGTTTTTGTGTTATCGTTTAAAGTGGCGAGGTTTAGG 960  
Qy 2236 AAGAGGGTCATTTCTTAAGCCACACATAGTCTGCACTCGTGGCTGCAGCCAAACAAAGA 2295  
Db 961 AAGAGGGTTATTTTAAAGTTATATATTTAGTTTGGTGTGTTAGTTTAAATTAAGA 1020  
Qy 2296 ACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGA 2355  
Db 1021 ATTGGGTGTGAGTATTTATTAATTAAGAATTAATAATTTAGGGTATTTATATGTGAAGGA 1080  
Qy 2356 TAAGAACTCATTCTCTTACTCTCCAAAGAGAGTGGGNAAGAACCAATCAACCTTTC 2415  
Db 1081 TAAGAAATTTATTTTATTTTATTTTAAAGAGAGTGGGNAAGAAATTAATAATTTT 1140  
Qy 2416 CTCCTGACTTACCAACAGCAAGAACAGCAGAGAGGTGGCTCAGGACTTAGGACAGG 2475  
Db 1141 TTTTGTGATTTATTAATTTAGGAAATATAGTAGAGAGGGTGTGTAGGATTTAGGGAATAGG 1200  
Qy 2476 GTATAGCTTAGATGGTGAAGCAAGAGAGAGGAGGAGTTGTAAATCACTGGCTAATGA 2535  
Db 1201 GTATAGTTTAGATGGTGAAGTAAGAGAGTAGGAGGTTGTAAATTTATGTTAATGA 1260  
Qy 2536 GAAAGAGACAGCTAATCTAGATGAGAGTGTGACTAGCTGGAGTGTGCTCCCTTGAA 2595  
Db 1261 GAAAGAGAGATGATTAATTTTAGATGAAGTGTGTAGGTGGAGTTGTGTTTTTGA 1320  
Qy 2596 GATGGGACTCTCGGTATCAAGACCTATGCCATCACACTGGGGCTAGGGAAGTAGGT 2655  
Db 1321 GATGGGATTTTGGGTATTAAGATTTATGTTATATATATGTTGGGTTAGGGAAGTAGGT 1380  
Qy 2656 GATCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTTTGAGAAGTTATATTTGGGAGTGGC 2715  
Db 1381 GATGTAGTTTAAAGTTGTTTTTGTAGTTAGGATTTGAGAAGTTATATTTGGGTAGTGT 1440  
Qy 2716 TCCAACTGTGGACAGATTTTACGCTTCCCTGAAGATCAGGAGGGTGCATCATG 2775  
Db 1441 TTTAAATTTGTGATTTAGTATTTTATTTTGTGAGATTTAGGTAGGTTGTTATTTATG 1500  
Qy 2776 TCTTTCTCTCTAGCCCTCAGGAAGAAGGACTATATTTGTACTGTACCTAGGGGTT 2835  
Db 1501 TTTTGTGTTTTTGTGTTTTTGTAGGAAGAAGATTTATTTGTATTTATTTAGGGGTT 1560  
Qy 2836 CTGGAAGGGAACATGGAATCAGGATTTATAGACTGATAGGCCCTATCCACAAGGCC 2895  
Db 1561 TTGGAAGGGAATAATGGAATTAGAATTTATAGATTGATGATGTTTATTTATAGGGTT 1620  
Qy 2896 ATGACTGGGAAGGATATGGAGCAGAGGAATTTGGGATTTTGGGTGAGCTACGCT 2955  
Db 1621 ATGATTTGGGAAGGATATGGGATAGAGGAGAAATTTGGGATTTTGGGTGTAGTTACGTT 1680  
Qy 2956 CACCTTAACCTTTTGGTGGCTGGGCAATGCTTGGGCCAGACTGTTAAGCAGGCTCT 3015  
Db 1681 TATTTTAAATTTTGGTGGTTTGGGGTATGTTTTGAGGTTTATGATTTGTTAATAGGTTTT 1740  
Qy 3016 GCTGGCTGTTTACTCGTCACCACTCTGCACTGCTGTCTTGGAGCTTCCATCCAGCCCC 3075  
Db 1741 GTTGGTTTGTATTCGTTATATTTTGTATTTTGTGTTTGTGATTTATTTAGTTTTT 1800  
Qy 3076 AGGCACGCCACTGCTCGTAGCCCTCCACTATCTCCCTGTGACCGGGTGAATCTCGTGTAC 3135  
Db 1801 AGGTACGTTATTTGTTTGTGTTTTTATTTATTTTGTGACGGGTGAATTTCTCGTGTAT 1860

Qy 3136 TGTGTCCTCGGTCATATATGAATTTGTAGCAGGGTTTCATCTATTTTAAACACAGATGTT 3195  
Db 1861 TGTGTCCTCGGTTTATATATGAATTTGTAGTAGGGTTTATTTATTTTAAATATAGATGTT 1920  
Qy 3196 TACAAATAAAGATTTATTTCAAAACACC 3223  
Db 1921 TATAAATAAAGATTTATTTTAAATATATC 1948  
  
RESULT 6  
AX347041/c  
LOCUS AX347041 7924 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 2112 from Patent WO0200928.  
ACCESSION AX347041  
VERSION AX347041.1 GI:18494929  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 2112 03-JAN-2002;  
Epigenomics AG (DE)  
FEATURES  
Location/Qualifiers  
source 1..7924  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
  
ORIGIN

Query Match 35.68; Score 1160.8; DB 6; Length 7924;  
Best Local Similarity 74.18; Pred. No. 1.7e-279; Indels 0; Gaps 0;  
Matches 1468; Conservative 0; Mismatches 512;  
  
Qy 1278 TGTCTTCTCTCTTCCGCCACCCCACTGTCCTCATGTGCCATGGCGCTGCCCTCCCGAG 1337  
Db 7922 TACTTCTCTCTTCCGCCACCCCACTGTCCTCATATACCATTAACCTACCTCCCGAA 7863  
Qy 1338 TGACCTCGGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGC 1397  
Db 7862 TAACCTACGAAATATAACATCGAAATATAAATAAAGCAACCAAAATATCTCGAAC 7803  
Qy 1398 CTGGGGTGGCTTACCTTACCCATTCGCCGACAGAGCTTGGCTTGTGCTGCC 1457  
Db 7802 CTAAACTACCTTACCTTACCCATTCGCCGACCAAACTTTACCTTACTTAACTACC 7743  
Qy 1458 GCCTGCTCTTTGGGAACTGAGCTCAGAGCAGGTGCTTCAGAGAGGAAACAAATGA 1517  
Db 7742 GCCTTACCTTTTAAATACTTAACCTCAAAACAATACTTCAAAAAATAAATAA 7683  
Qy 1518 GGGGTGCGAGGATATAAAGTCACTCTCATCTCTCTCTCCATGCGAGCATGAACAT 1577  
Db 7682 AAAATAACAAAAATAAATAATCACCTCAATCTCTACCTCCATACCAACATATAACA 7623  
Qy 1578 TTCTCTCAGCTGGCTCCCAATTTAAAGATGTGGACCAAGCCCTGTGGGTACTCCAGG 1637  
Db 7622 TTCTCTCCCTTACCTCCCAATTTAAATAATATAAACAACCACTTAAATACTTCAAAA 7563  
Qy 1638 GCAAGGAGAGCCCTGGGCTCAGTGACACTGTGAGGCCAACCATGCATCTCCCAAGGGGA 1697  
Db 7562 ACAAAAAAACCTTAATCAATCACTATCAAAACCAACATCACTCCCAAAAAA 7503  
Qy 1698 GCATTTGGAATGAAGGACTAGCTCTTATGTATCAGGTTTAAGAGCAAGGGAGCTGCC 1757  
Db 7502 ACATTTAAAAATAAATAAATAATCACTCTATATCAATTTAAATAAATAAATAAATAA 7443  
Qy 1758 AGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTTAGGCCCA 1817  
Db 7442 AAAACCAACATTTTACCAACAAAAATAATATAAACAACCAAAACCTCTCTAAACCCCA 7383







Best Local Similarity 99.6%; Pred. No. 2.3e-239; Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	2	GCCGCTGCCGCGCCGAGTTC	CCGCGCCGCTGCGCCGAGTACG	CGGAGGAGTACGA	61				
Db	38	GCCGCTGCCGCGCCGAGTTC	CCGCGCCGCTGCGCCGAGTACG	CGGAGGAGTACGA	97				
QY	62	TGTCCTGTCGGCTGCTGCT	GATCGGGGACTCCGGGGTGGG	CAAGACCTGCGCTGCTG	121				
Db	98	TGTCCTGTCGGCTGCTGCT	GATCGGGGACTCCGGGGTGGG	CAAGACCTGCGCTGCTG	157				
QY	122	CCGCTTCAACGACAAAGGTT	CCACTCCTCGCACATCTCC	ACCATCGGTGTGACTTTAA	181				
Db	158	CCGCTTCAACGACAAAGGTT	CCACTCCTCGCACATCTCC	ACCATCGGTGTGACTTTAA	217				
QY	182	GATGAGACCATAGAGTAG	CGGCATCAAGTGGGATAC	GATCTCGGACACTGACG	241				
Db	218	GATGAGACCATAGAGTAG	CGGCATCAAGTGGGATAC	GATCTCGGACACTGACG	277				
QY	242	GCAGGAGATACAGACCAT	CAAAAGCAGTACTATCGG	CGGCCCCAGGGGATATTTT	301				
Db	278	GCAGGAGATACAGACCAT	CAAAAGCAGTACTATCGG	CGGCCCCAGGGGATATTTT	337				
QY	302	GGTCTATGACATTAGCAG	CGCCTCTTACAGCACAT	CATGAAATGGGTGAGTACGT	361				
Db	338	GGTCTATGACATTAGCAG	CGCCTCTTACAGCACAT	CATGAAATGGGTGAGTACGT	397				
QY	362	GGATGATGCGCACCAAG	CGCTCCAGAGATCTTAT	TGGGATGAGTGTATGAGGA	421				
Db	398	GGATGATGCGCACCAAG	CGCTCCAGAGATCTTAT	TGGGATGAGTGTATGAGGA	457				
QY	422	GCAGAAACGCGCAGTGG	GAAGAGAGCAGCAGCTG	CGGAGGAGTATGGCATGGA	481				
Db	458	GCAGAAACGCGCAGTGG	GAAGAGAGCAGCAGCTG	CGGAGGAGTATGGCATGGA	517				
QY	482	CTTCTATGAAACAGTGC	CTGCACCACTCAACATTA	AGAGTCAITTCACGCTCTGAC	541				
Db	518	CTTCTATGAAACAGTGC	CTGCACCACTCAACATTA	AGAGTCAITTCACGCTCTGAC	577				
QY	542	AGAGCTGGTCTCGAGCC	CATAGGAGAGCTGGA	AGGCTCCGATGCTGCAGCAA	601				
Db	578	AGAGCTGGTCTCGAGCC	CATAGGAGAGCTGGA	AGGCTCCGATGCTGCAGCAA	637				
QY	602	TGAGTTGGCACTGGCAG	AGCTGGAGGAGGAGG	CAAAACCCAGGCGCCAGCAACTC	661				
Db	638	TGAGTTGGCACTGGCAG	AGCTGGAGGAGGAGG	CAAAACCCAGGCGCCAGCAACTC	697				
QY	662	TTGGAACCTGCTGGTGA	GTGCTGTGTGGGGCA	ACCCACACACCCCTCTTCCC	721				
Db	698	TTGGAACCTGCTGGTGA	GTGCTGTGTGGGGCA	ACCCACACACCCCTCTTCCC	757				
QY	722	TCAGGAGCCGCTGGGCA	GACGGGAGCGGGCTT	TGCGCTGCTCTCTGCTG	781				
Db	758	TCAGGAGCCGCTGGGCA	GACGGGAGCGGGCTT	TGCGCTGCTCTCTGCTG	817				
QY	782	TGATGACCTTATGATAT	CAGTAGCCACTACTCC	CGCTGCTGCTGCTGCTG	841				
Db	818	TGATGACCTTATGATAT	CAGTAGCCACTACTCC	CGCTGCTGCTGCTGCTG	877				
QY	842	TGCTGTCACTCAAGCAG	CCCTGTGCTCCAGCCG	CTCCACCTGAGTGTCTTTCAG	901				
Db	878	TGCTGTCACTCAAGCAG	CCCTGTGCTCCAGCCG	CTCCACCTGAGTGTCTTTCAG	937				
QY	902	CTGTGTTTCCCGCCAG	CCAGGCTGTACACCC	CCGAGTGTCCGCAAGCACTGTCTCA	961				
Db	938	CTGTGTTTCCCGCCAG	CCAGGCTGTACACCC	CCGAGTGTCCGCAAGCACTGTCTCA	997				
QY	962	CCATCCCGCACCCAC	CAAGACAGCGGCTG	GAGTCCAGGCCA 1007					
Db	998	CCATCCCGCACCCAC	CAAGACAGCGGCTG	GAGTCCAGGCCA 1043					
.RESULT 9									

AX014147	LOCUS	AX014147	Sequence 12 from Patent WO9954447.	895 bp	DNA	linear	PAT 07-SEP-2000
AX014147	ACCESSION	AX014147					
AX014147.1	VERSION	AX014147.1	GI:10040594				
KEYWORDS	SOURCE						
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
ORIGIN							
Query Match							
Best Local Similarity							
Matches 879; Conservative							
0; Mismatches 3; Indels 1; Gaps 1;							
QY	2358	AGAACCTCACTTCTCTCT	CTCTCAAAAAGAGTGGG	AAAGAACCATCAAACTTCTCT	2417		
Db	14	AGAACCTCACTTCTCTCT	CTCTCAAAAAGAGTGGG	AAAGAACCATCAAACTTCTCT	73		
QY	2418	CCTGACTTACCAACAC	GAGAAACACGAGAGG	TGGCTCAGGACTTAGGACAGGCT	2477		
Db	74	CCTGACTTACCAACAC	GAGAAACACGAGAGG	TGGCTCAGGACTTAGGACAGGCT	133		
QY	2478	ATAGCTTAGATGCTGA	AGAAAGAGAGAGTGT	TAATCACTGCTTAATGAGA	2537		
Db	134	ATAGCTTAGATGCTGA	AGAAAGAGAGAGTGT	TAATCACTGCTTAATGAGA	193		
QY	2538	AAAGGAGACAGCTAA	CTCTAGGATGAAGCTG	TGCTAGGCTGGAGTTGCTTCTTGAAGA	2597		
Db	194	AAAGGAGACAGCTAA	CTCTAGGATGAAGCTG	TGCTAGGCTGGAGTTGCTTCTTGAAGA	253		
QY	2598	TGGGACTCTTGGGTAT	CAAGACCTATGCCA	TATCACACTGGGGCTAGGGAAGTAGTGA	2657		
Db	254	TGGGACTCTTGGGTAT	CAAGACCTATGCCA	TATCACACTGGGGCTAGGGAAGTAGTGA	313		
QY	2658	TGCCAGCCCTCAAGT	CTGTCTTCAGCCAGG	ACTTCAGAAAGTTATTTGGGCAAGTGGCTC	2717		
Db	314	TGCCAGCCCTCAAGT	CTGTCTTCAGCCAGG	ACTTCAGAAAGTTATTTGGGCAAGTGGCTC	373		
QY	2718	CAATCTGTGGACCA	GATTTTTCAGCTTTC	CTGAAGATCAGGCAAGGCTGCGCATTTCAITGTC	2777		
Db	374	CAATCTGTGGACCA	GATTTTTCAGCTTTC	CTGAAGATCAGGCAAGGCTGCGCATTTCAITGTC	433		
QY	2778	TTTCTCTCTAGCC	CTCAGGAAAGAGCT	TATTTTGTACTGTACCTCAGGGGTTCT	2837		
Db	434	TTTCTCTCTAGCC	CTCAGGAAAGAGCT	TATTTTGTACTGTACCTCAGGGGTTCT	493		
QY	2838	GGAAGGAAAAA	CATGGAATCAGG	ATTTCTATAGCTATAGGCCCTTATCCACAAGGGCCAT	2897		
Db	494	GGAAGGAAAAA	CATGGAATCAGG	ATTTCTATAGCTATAGGCCCTTATCCACAAGGGCCAT	553		
QY	2898	GACTGGAAAAAG	GTATGGGAGCAAG	AGAAATTTGGGATTTTATAGGCTGACGCTACGCTCA	2957		
Db	554	GACTGGAAAAAG	GTATGGGAGCAAG	AGAAATTTGGGATTTTATAGGCTGACG - TACGCTCA	612		
QY	2958	CCCTAACTTTTGG	TGGCTTGGGCAAT	GTCTTGGAGCCCAAGACTGTTAAGCAGGCTCTGC	3017		
Db	613	CCCTAACTTTTGG	TGGCTTGGGCAAT	GTCTTGGAGCCCAAGACTGTTAAGCAGGCTCTGC	672		



Qy	3018	TGGCTGTTTACTCGTCACCACTCTGACACCTGCTGCTTGTGAGACTCCATCCAGCCCGAG	3077
Db	673	TGGCTGTTTACTCGTCACCACTCTGACACCTGCTGCTTGTGAGACTCCATCCAGCCCGAG	732
Qy	3078	GCAGCCACTGCTCCTGAGCCTCCACTATCTCCCTGTGAGCGGTGAACCTCGTGTACTG	3137
Db	733	GCAGCCACTGCTCCTGAGCCTCCACTATCTCCCTGTGAGCGGTGAACCTCGTGTACTG	792
Qy	3138	TGCTCTGGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTA	3197
Db	793	TGCTCTGGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTA	852
Qy	3198	CAAAATAAAGATTATTTCAACACCAACCAAAAAA	3240
Db	853	CAAAATAAAGATTATTTCAACACCAACCAAAAAA	895
RESULT 10			
BD205055		895 bp DNA linear	PAT 17-JUL-2003
LOCUS			
DEFINITION		Human nucleic acid sequence originating in cystic cancer tissue.	
ACCESSION		BD205055	
VERSION		BD205055.1 GI:33014825	
KEYWORDS		JP 2002512023-A/9.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE			
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Specht, T., Hinzmann, B., Schmitt, A., Pilarsky, C., Dahl, E. and Rosenthal, A.	
TITLE		Human nucleic acid sequence originating in cystic cancer tissue	
JOURNAL		Patent: JP 2002512023-A 9 23-APR-2002;	
COMMENT		METAGEN GESELTSCHAFT FUER GENOM FORSCHUNG MBH	
		OS Homo sapiens (human)	
		PN JP 2002512023-A/9	
		PD 23-APR-2002	
		PP 15-APR-1999 JP 2000544779	
		PR 21-APR-1998 DE 198 18 619, 3	
		PI THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY, PI EDGAR DAHL, PI ANDRE ROSENTHAL	
		PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P13/10, A61P35/00,	
		PC C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12P21/08, C12Q1/68, PC C12N15/00,	
		PC A61K37/02, C12N5/00	
		CC Human nucleic acid sequence originating in cystic cancer tissue	
FEATURES			
source			
		FT Key Location/Qualifiers	
		FT source 1..895 /organism='Homo sapiens (human)'	
		1..895 Location/Qualifiers	
		/organism='Homo sapiens'	
		/mol_type='genomic DNA'	
		/db_xref='taxon:9606'	
ORIGIN			
Query Match		26.6%; Score 866.2; DB 6; Length 895;	
Best Local Similarity		99.5%; Pred. No. 8.9e-206;	
Matches		879; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	
Qy	2358	AGAACCTCCTCTTCTCTCTCCAAAAGAGTGGGAAAGAACCATCAACCTTTCT	2417
Db	14	AGAACCTCCTCTTCTCTCTCCAAAAGAGTGGGAAAGAACCATCAACCTTTCT	73
Qy	2418	CCTGACTTACCAACAGGAAACAGCAGGAGGTGCTCAGGACTTAGGACAGGT	2477
Db	74	CCTGACTTACCAACAGGAAACAGCAGGAGGTGCTCAGGACTTAGGACAGGT	133
Qy	2478	ATAGCTTAGTGGTGGAAAGCAAGGAGACGAGGAGGTGTTAAATCACTGGCTATGAGA	2537

Db	134	ATAGCTTAGATGTGGAAAGCAAGGAGACAGGAAGTTGTAATCACTGGCTAATGAGA	193
Qy	2538	AAAGGACACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTTGAAGA	2597
Db	194	AAAGGACACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTTGAAGA	253
Qy	2598	TGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGA	2657
Db	254	TGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGA	313
Qy	2658	TGCAGCCCTCAAGTCTGCTTCCAGCCAGGACTTTGAGAAAGTTATATTGGGACGTGGCTC	2717
Db	314	TGCAGCCCTCAAGTCTGCTTCCAGCCAGGACTTTGAGAAAGTTATATTGGGACGTGGCTC	373
Qy	2718	CAATCTGTGACACCATGATTTTTCAGCTTTCCCTCAAGATCAGCAGGGTGCCATTCAATTGTC	2777
Db	374	CAATCTGTGACACCATGATTTTTCAGCTTTCCCTCAAGATCAGCAGGGTGCCATTCAATTGTC	433
Qy	2778	TTTCTCTCTAGCCCTCAGGAAAGGACTATATTGTTACTGTGTACCTTAGGGTTCT	2837
Db	434	TTTCTCTCTAGCCCTCAGGAAAGGACTATATTGTTACTGTGTACCTTAGGGTTCT	493
Qy	2838	GGAAGGAAACATGGAATCAGGATTTATAGACTGATAGCCCTATCCCAAGGGCCAT	2897
Db	494	GGAAGGAAACATGGAATCAGGATTTATAGACTGATAGCCCTATCCCAAGGGCCAT	553
Qy	2898	GACTGGGAAAGGTATGGGAGCAGAGGAGATTTGGGATTTTAGGCTGCAGCTACGCTCA	2957
Db	554	GACTGGGAAAGGTATGGGAGCAGAGGAGATTTGGGATTTTAGGCTGCAG-TACGCTCA	612
Qy	2958	CCCTAAACTTTTGGTGGCCTGGGCACTGTCTTGAGGCCAGACTGTTAAAGCAGGCTCTGC	3017
Db	613	CCCTAAACTTTTGGTGGCCTGGGCACTGTCTTGAGGCCAGACTGTTAAAGCAGGCTCTGC	672
Qy	3018	TGGCTGTTTACTGCTCACACCTCTGCACCTGCTGCTTGTGAGACTCCATCCAGCCCGAG	3077
Db	673	TGGCTGTTTACTGCTCACACCTCTGCACCTGCTGCTTGTGAGACTCCATCCAGCCCGAG	732
Qy	3078	GCAGCCACTGCTCTGAGCCTCCACTATCTCCCTGTGAGCGGTGAACCTTCGTGTACTG	3137
Db	733	GCAGCCACTGCTCTGAGCCTCCACTATCTCCCTGTGAGCGGTGAACCTTCGTGTACTG	792
Qy	3138	TGCTCTGGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTA	3197
Db	793	TGCTCTGGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTA	852
Qy	3198	CAAAATAAAGATTATTTCAACACCAACCAAAAAA	3240
Db	853	CAAAATAAAGATTATTTCAACACCAACCAAAAAA	895
RESULT 11			
LOCUS			
DEFINITION		Mus musculus RAB15, member RAS oncogene family, mRNA (cDNA clone MGC:38375 IMAGE:5345297), complete cds.	
ACCESSION		BC027769	
VERSION		BC027769.1 GI:20380721	
KEYWORDS		MGC.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE			
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3139)	
		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,	





181 GATCAGCGGGTGCATTCTGCTCTCTCTAGCCCTCAGGAAGAAGACTA 240  
2812 TATTGTACTGTACCTAGGGGTTCTGGAGGGAACATGGATCAGGATCTTAGAC 2871  
241 TATTGTACTGTACCTAGGGGTTCTGGAGGGAACATGGATCAGGATCTTAGAC 300  
2872 TGATAGCCCTATCCCAAGGGCCATGACCTGGGAAAGGTATGGAGCAAGAGAAATT 2931  
301 TGATAGCCCTATCCCAAGGGCCATGACCTGGGAAAGGTATGGAGCAAGAGAAATT 360  
2932 GGGATTTTGGGTGCAGTACGCTCACCTAAACCTTTTGGTGGGCTGGGCACTGTCTTGA 2991  
361 GGGATTTTGGGTGCAGTACGCTCACCTAAACCTTTTGGTGGGCTGGGCACTGTCTTGA 420  
2992 GGCCAGACGTGTTAAGCAGCTCTGCTGGGCTGTTTACTCGTCAACCTCTGCACTGTC 3051  
421 GGCCAGACGTGTTAAGCAGCTCTGCTGGGCTGTTTACTCGTCAACCTCTGCACTGTC 480  
3052 TGTCTTGAGACTCCATCCAGCCCGCAGGACCGCCTGCTCCCTGAGCCTCCACTATCTCC 3111  
481 TGTCTTGAGACTCCANN 540  
3112 CTGTGACGGGTGAATCTCGTGTACTGTCTCGGGTCCATATATGAATTTGAGCAGGGT 3171  
541 CTGTGACGGGTGAATCTCGTGTACTGTCTCGGGTCCATATATGAATTTGAGCAGGGT 600  
3172 TCATCTATTTTAAACACAGATGTTTACAAATAAAGATTTTCAAACCCACCAAAAAA 3231  
601 TCATCTATTTTAAACACAGATGTTTACAAATAAAGATTTTCAAACCCACCAAAAAA 660  
3232 AAAAAAAGAAAAA 3247  
661 AAAAAAAGAAAAA 676

RESULT 13  
PATRAB15X  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
FEATURES  
source  
1. 945  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/tissue\_lib="LAMBDA ZAPII"  
1. 945  
/gene="RAB15"  
220. 858  
/gene="RAB15"  
/codon\_start=1  
/product="RAB15"  
/protein\_id="AAA41995.1"

RATRAB15X 945 bp mRNA linear ROD 27-APR-1993  
Sprague-Dawley (clone LRB9), RAB15 mRNA, complete cds.  
M83679.1 GI:206536  
LMW GTP-binding protein.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 945)  
Elferink, L.A., Anzai, K. and Scheller, R.H.  
Rab15, a novel, low molecular weight GTP-binding protein  
specifically expressed in rat brain  
J. Biol. Chem. 267 (9), 5768-5775 (1992)  
1313420  
Original source text: Rattus norvegicus (strain Sprague-Dawley)  
(library: LAMBDA ZAPII) adult brain cDNA to mRNA.  
Location/Qualifiers  
1. 945  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/tissue\_lib="LAMBDA ZAPII"  
1. 945  
/gene="RAB15"  
220. 858  
/gene="RAB15"  
/codon\_start=1  
/product="RAB15"  
/protein\_id="AAA41995.1"

/db\_xref="GI:206537"  
/translation="MAKQYDVLFRLLILGDSGVGKTKLLCRFTDNFHHSHSTIGVD  
FMKTIEDVGIKVIQIWDTAGERYQITIKYVRRQAQIFLYVYDISSERSYQHIMKW  
VSDVBYAPEGVKILIGNKADEBOKROVGROGOLAKKEYGMDPVEYSACTNLNIKE  
SFTRLITELVLOHRRKELDLRLTCASNELALAELEDEGKTEGPANSSKTCWC"  
ORIGIN  
Query Match 17.7%; Score 577.6; DB 10; Length 945;  
Best Local Similarity 88.7%; Pred. No. 2.1e-133;  
Matches 637; Conservative 0; Mismatches 79; Indels 2; Gaps 1;  
QY 21 TTCCCGGCCCGCTGGCCCGAGTCATGGCGAAGCAGTACGATGTGTCTTCGGCTGCTG 80  
Db 196 TTCCCGGCCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255  
QY 81 CTGATCGGGGATCCCGGGTGGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140  
Db 256 CTGATCGGGGATCCCGGGTGGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315  
QY 141 TTCCACTCTCGCACATCTCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 200  
Db 316 TTCCACTCTCGCATATCTCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375  
QY 201 GACGGCATCAAGTGGGATACAGATCTGGGACACATCTGCGGGGAGGAGGATACCGAGC 260  
Db 376 GACGGCATCAAGTGGGATACAGATCTGGGACACATCTGCGGGGAGGAGGATACCGAGC 435  
QY 261 ATCACAAGCAGTACTCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 320  
Db 436 ATCACAAGCAGTACTCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 495  
QY 321 GAGCGCTCTTACCAGCACATCATGAAAGTGGGTGAGTGGATGAGTACGACACGACAGAA 380  
Db 496 GAGCGCTCTTACCAGCACATCATGAAAGTGGGTGAGTGGATGAGTACGACACGACAGAA 555  
QY 381 GCGCTCAGAAAGATCTTATTTGGGATTAAGCTGTGATGAGGAGGAGGAGGAGGAGGAGG 440  
Db 556 GCGCTCAGAAAGATCTTATTTGGGATTAAGCTGTGATGAGGAGGAGGAGGAGGAGGAGG 615  
QY 441 AGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 500  
Db 616 AGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 675  
QY 501 TGCACCAACCTCAACATTTAAAGAGTCACTCACGCTGTGACAGAGCTGTGCTGCGAGGCC 560  
Db 676 TGCACCAACCTCAACATTTAAAGAGTCTTCACTCGTGTGACCGAGCTGTGCTGCGAGGC 735  
QY 561 CATAGGAAGAGCTGGAAGGCTCCGGATCGGTGCCAGCAATGAGTTGGCACTGGCAGAG 620  
Db 736 CACAGGAAGAGCTGGAATGCTCTCCGAACATGTGCCAGCAATGAGTCTGCACTGGCCGAG 795  
QY 621 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 680  
Db 796 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 855  
QY 681 TGA--GTCCTGTGTGGGGCACCCACACGACACCCCTCTTCCCTCAGGAGGCCGCTGG 736  
Db 856 TGAGGGTCTGTGTGGGAATCTCTCCCTCCCGAGTCCCTTCCCTCAGGAGGCCGCTGG 913  
RESULT 14  
AX396088  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
AX396088 Sequence 303 from Patent WO0212328.  
AX396088  
AX396088.1 GI:21066835  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
King, G.E., Meagher, M.J., Xu, J. and Secrist, H.

TITLE Compositions and methods for the therapy and diagnosis of colon cancer  
JOURNAL Patent: WO 0212328-A 303 14-FEB-2002;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..481  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 14.7%; Score 477.4; DB 6; Length 481;  
Best Local Similarity 99.4%; Pred. No. 2.6e-108;  
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2582 GTTCTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2641  
Db 1 GTTCTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 60  
QY 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGACTTGAAGTTA 2701  
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGACTTGAAGTTA 120  
QY 2702 TATTGGGAGTGGTCCAAATCTGGACCAAGTATTTAGCTTTCCCTGAAGATCAGGCAG 2761  
Db 121 TATTGGGAGTGGTCCAAATCTGGACCAAGTATTTAGCTTTCCCTGAAGATCAGGCAG 180  
QY 2762 GGTGCCATTCAATCTCTTCTCTCTAGCCCTCAGGAAGAGGACTATATTGTACT 2821  
Db 181 GGTGCCATTCAATCTCTCTCTAGCCCTCAGGAAGAGGACTATATTGTACT 240  
QY 2822 GTACCTAGGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2881  
Db 241 GTACCTAGGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 300  
QY 2882 TATCCACAAGGCCATGACTGGGAAAAAGTATGGAGCAGAGAAATTTGGGATTTTGG 2941  
Db 301 TATCCACAAGGCCATGACTGGGAAAAAGTATGGAGCAGAGAAATTTGGGATTTTGG 360  
QY 2942 GGTGACGTACGCTCACCCTAACTTTTGGTGGCTGGGCTCTTGGAGCCAGACT 3001  
Db 361 GGTGACGTACGCTCACCCTAACTTTTGGTGGCTGGGCTCTTGGAGCCAGACT 420  
QY 3002 GTTAAGCAGGCTCTGTGGCTGTTTACTCTGCTACCACTCTGACCTGCTGTTGAGA 3061  
Db 421 GTTAACAGGCTCTGTGGCTGTTTACTCTGCTACCACTCTGACCTGTTGAGA 480  
QY 3062 C 3062  
Db 481 C 481

RESULT 15  
AX334820/c  
LOCUS AX334820 463 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 5329 from Patent WO0194629.  
ACCESSION AX334820  
VERSION AX334820.1 GI:18125539  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrikan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 5329 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES Location/Qualifiers  
source 1..463  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 13.5%; Score 438.8; DB 6; Length 463;  
Best Local Similarity 99.3%; Pred. No. 1.2e-96;  
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 2790 CCCCTCAGGAAGAAGGACTATATTGTACTGTACCTAGGGGTCTTGGGAAGGAAAAAC 2849  
Db 453 CCCCTCAGGAAGAAGGACTATATTGTACTGTACCTAGGGGTCTTGGGAAGGAAAAAC 394  
QY 2850 ATGAATCAGGATTTCTATAGACTATAGGCCCTATCCACAAGGCCCATGACTGGGAAAAAG 2909  
Db 393 ATGAATCAGGATTTCTATAGACTATAGGCCCTATCCACAAGGCCCATGACTGGGAAAAAG 334  
QY 2910 GTATGGGAGCAGAGGAGAAATTTGGGATTTAGGGTGCAGCTACCTCACCCCTAACTTTT 2969  
Db 333 GTATGGGAGCAGAGGAGAAATTTGGGATTTAGGGTGCAG-TACGCTCACCCCTAACTTTT 275  
QY 2970 GGTGGCCTGGGGCATGTCTTGAGGCCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 3029  
Db 274 GGTGGCCTGGGGTATGTCTTGAGGCCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 215  
QY 3030 TCGTCACCACTCTGCACCTGCTGTCTTGAGACTTCCATCCAGCCCCCAGGCCACCTG 3089  
Db 214 TCGTCACCACTCTGCACCTGCTGTCTTGAGACTTCCATCCAGCCCCCAGGCCACCTG 155  
QY 3090 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCC 3149  
Db 154 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCC 95  
QY 3150 ATATATGAATTTGTGAGCAGGCTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 3209  
Db 94 ATATATGAATTTGTGAGCAGGCTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 35  
QY 3210 TATTTCAAAACCCAAAAAATAAAAAA 3243  
Db 34 TATTTCAAAACCCAAAAAATAAAAAA 1

Search completed: March 21, 2004, 08:59:33  
Job time : 12136 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 02:17:36 ; Search time 1147 Seconds  
(without alignments)  
12063.105 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcccgctgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002s:\*

7: Geneseq2003as:\*

8: Geneseq2003bs:\*

9: Geneseq2003cs:\*

10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3257	100.0	3257	7	AAD47168	Aad47168 Human Ras
2	2696.6	82.8	3240	4	AAK82285	Aak82285 Human imm
3	2696.6	82.8	28770	7	AAD47169	Aad47169 Human Ras
4	1568.2	48.1	1666	3	AAK75813	Aac75813 Human ORF
5	1192.8	36.6	7924	6	ABK40069	Abk40069 Human che
6	1192.8	36.6	7924	6	ABL34138	Abi34138 Human imm
7	1160.8	35.6	7924	6	ABK40070	Abk40070 Human che
8	1160.8	35.6	7924	6	ABL34139	Abi34139 Human imm
9	1160.6	35.6	2021	4	AAK27053	Aas27053 CDNA enco
10	1160.6	35.6	2021	9	ABK93231	Abk93231 Human cDN
11	1000.2	30.7	1054	6	ABK93361	Abq93361 Human cDN
12	866.2	26.6	895	2	AAZ24400	Aaz24400 Human bla
13	534.2	16.4	566	4	AAK27472	Aas27472 CDNA enco
14	534.2	16.4	566	9	ABK93650	Abk93650 Human cDN
15	477.4	14.7	481	6	ABK44752	Abk44752 CDNA enco
16	438.8	13.5	463	6	ABK66992	Abi66992 Thyroid c
17	424.4	13.0	458	8	ACH33636	Ach33636 Human end
18	313.4	9.6	320	2	AAT25457	Aat25457 Human gen
19	307	9.4	313	6	ABL38341	Abi38341 Human col
20	229	7.0	1986	4	ABQ09160	Abq09160 Human rab
21	228.4	7.0	2411	6	ABQ54410	Abq54410 Human ova
22	217.8	6.7	1161	4	AAH75182	Aah75182 Nucleotid
23	217.8	6.7	1265	6	ABV78060	Abv78060 Hypoxia-r

24	217.8	6.7	2497	4	AAS60878	Aas60878 Human can
25	217.8	6.7	2497	4	AAS60884	Aas60884 Human can
26	217.8	6.7	2497	4	AAS60893	Aas60893 Human can
27	217.8	6.7	2497	4	AAS60895	Aas60895 Human can
28	217.8	6.7	2497	5	ABV25781	Abv25781 Human pro
29	217.8	6.7	2497	5	ABV30037	Abv30037 Human pro
30	217.8	6.7	3077	4	AAH13912	Aah13912 Human cDN
31	214.2	6.6	866	4	AAH04301	Aah04301 Human cDN
32	209	6.4	1859	8	ADB12721	Adb12721 Human pol
33	203.2	6.2	991	9	ADB52880	Adb52880 Primary r
34	202.6	6.2	412	3	AAK01700	Aac01700 Human sec
35	200.2	6.1	716	3	AAA40104	Aaa40104 Human Rab
36	200.2	6.1	861	3	AAA40108	Aaa40108 Human Rab
37	198.4	6.1	956	3	AAAG6887	Aaa96887 Nucleotid
38	198.4	6.1	1537	3	AAF16196	Aaf16196 Human pro
39	198.4	6.1	3533	4	AAH17889	Aah17889 Human cDN
40	196.8	6.0	674	5	AAH71453	Aas71453 Human enco
41	196.6	6.0	2247	4	ABL29661	Abi29661 Drosophil
42	195.6	6.0	3164	6	ABX04179	Abx04179 Human mRN
43	195.2	6.0	888	3	AAA40109	Aaa40109 Canine Ra
44	192	5.9	1274	3	AAF18072	Aaf18072 Lung canc
45	188.6	5.8	893	5	AAS87552	Aas87552 DNA enco

ALIGNMENTS

RESULT 1  
RAD47168  
ID AAD47168 standard; cDNA; 3257 BP.  
XX  
AC AAD47168;  
XX  
XX  
DT 24-FEB-2003 (first entry)  
XX  
DE Human Ras-like protein encoding cDNA.  
XX  
KW Human; Ras-like protein; inflammatory; cell proliferation; apoptosis;  
KW immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
KW Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
KW osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
KW irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
KW pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
KW Sjogren's syndrome; infection; transgenic; gene therapy; neotropic; gout;  
KW neuroprotective; vasotropic; cytostatic; dermatological; nephrotropic;  
KW ophthalmological; tranquiliser; cancer; stroke; Grave's disease; AIDS;  
KW asthma; anaemia; drug screening; gene; ss.  
XX  
OS Homo sapiens.

Key	Location/Qualifiers
5'UTR	1..44
CDS	/*tag= a 45..683 /*tag= b
3'UTR	/product= "Human Ras-like protein" 684..3257 /*tag= c

WO200277193-A2.

03-OCT-2002.

27-MAR-2002; 2002WO-US009328.

27-MAR-2001; 2001US-00817198.

(PEKE ) PE CORP.

Gan W, Ye J, Di Francesco V, Beasley EM;

WPI; 2003-018913/01.

P-PSDB; AAE29096.

XX New isolated human Ras-like protein polypeptide, useful for diagnosing,  
PT treating or preventing inflammation and disorders associated with cell  
PT proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma  
PT or stroke.

XX Claim 4; Page 73-74; 82pp; English.

XX The invention relates to human Ras-like protein and its corresponding  
CC nucleic acid. The Ras-like protein and DNA is useful in the development  
CC of human therapeutics and diagnostic compositions. They are useful in the  
CC diagnosis, prevention and treatment of inflammation and disorders  
CC associated with cell proliferation and apoptosis, e.g. AIDS and other  
CC infectious or genetic immunodeficiencies, neurodegenerative disease e.g.  
CC Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,  
CC ischaemic injuries e.g. myocardial infarction, stroke or reperfusion  
CC injury, toxin-induced diseases such as alcohol-induced liver damage or  
CC cirrhosis, osteoporosis or cancer. They are also used to treat disorders  
CC associated with inflammation including allergies, atopic dermatitis,  
CC atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,  
CC Grave's disease, glomerulonephritis, gout, irritable bowel syndrome,  
CC lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,  
CC autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,  
CC uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or  
CC helminthic infections. The antibodies of the invention are useful in  
CC pharmacogenomic analysis or for tissue typing. The transgenic animals are  
CC useful for studying the function of a Ras-like protein, and identifying  
CC and evaluating modulators of its activity. Ras-like protein is used in  
CC drug screening assays and its DNA is used in gene therapy. The present  
CC sequence is human Ras-like protein encoding cDNA

XX Sequence 3257 BP; 793 A; 892 C; 870 G; 702 T; 0 U; 0 Other;

Query Match 100.0%; Score 3257; DB 7; Length 3257;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCCGCTGCGCCGCGAGTTCCCGGCCCGCTGGCCCGAGTCATCGCGAAGCAGTACG 60  
DB 1 TGCCCGCTGCGCCGCGAGTTCCCGGCCCGCTGGCCCGAGTCATCGCGAAGCAGTACG 60  
QY 61 ATGTCGTGTCGGCTGCTGATCGGGGACTCCGGGGTGGGCAAGCCTGCTGCTGT 120  
DB 61 ATGTCGTGTCGGCTGCTGATCGGGGACTCCGGGGTGGGCAAGCCTGCTGCTGT 120  
QY 121 GCCGCTTACCGCAAGAGTTCCACTCTCTCGACATCTCCACCCTGGTGTGACTTTA 180  
DB 121 GCCGCTTACCGCAAGAGTTCCACTCTCTCGACATCTCCACCCTGGTGTGACTTTA 180  
QY 181 AGATGAAGACCATAGAGTGAGCGGCATCAAGTGCGGATACAGATCTGGGACATCGCAG 240  
DB 181 AGATGAAGACCATAGAGTGAGCGGCATCAAGTGCGGATACAGATCTGGGACATCGCAG 240  
QY 241 GCGAGGAGATACCAAGCAGTACTATCGGGGGCCCGAGGAGTATTTT 300  
DB 241 GCGAGGAGATACCAAGCAGTACTATCGGGGGCCCGAGGAGTATTTT 300  
QY 301 TGGTCTATGATGATGAGCGGCTTACAGGACATCATGAACTGGGTGAGTACG 360  
DB 301 TGGTCTATGATGATGAGCGGCTTACAGGACATCATGAACTGGGTGAGTACG 360  
QY 361 TGGATGATGATGAGCGGCTTACAGGAGATCTTATTTGGGAATAAGGCTGATGAGG 420  
DB 361 TGGATGATGATGAGCGGCTTACAGGAGATCTTATTTGGGAATAAGGCTGATGAGG 420  
QY 421 AGCAGAAACGGCGAGTGGGAAGAGCAAGGCGAGCAGCTGGGCAAGGATATGGCATGG 480  
DB 421 AGCAGAAACGGCGAGTGGGAAGAGCAAGGCGAGCAGCTGGGCAAGGATATGGCATGG 480  
QY 481 ACTTCTATGAACAAGTGGCTGACCAACCTTACATTTAAGAGTCACTTACCGGCTCTGA 540  
DB 481 ACTTCTATGAACAAGTGGCTGACCAACCTTACATTTAAGAGTCACTTACCGGCTCTGA 540

QY 541 CAGAGCTGCTGCTCGAGGCCCATAGGAAGAGAGCTGGAAGGCTCCGGATGCTGCCAGCA 600  
DB 541 CAGAGCTGCTGCTCGAGGCCCATAGGAAGAGAGCTGGAAGGCTCCGGATGCTGCCAGCA 600  
QY 601 ATGAGTTGGCACTGGCGAGAGCTGGAGAGAGGCGCAAAACCCGAGGGGCCAGCAACT 660  
DB 601 ATGAGTTGGCACTGGCGAGAGCTGGAGAGAGGCGCAAAACCCGAGGGGCCAGCAACT 660  
QY 661 CTTTCGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 661 CTTTCGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 CTACGAGGAGGCGCTGGGCGACAGAGGAGGCGGGGCTTTGCGCTGCTGCTGCTGCTGCT 780  
DB 721 CTACGAGGAGGCGCTGGGCGACAGAGGAGGCGGGGCTTTGCGCTGCTGCTGCTGCTGCT 780  
QY 781 GTGATGACCTATTTAGTATFAGTACGATAGTACCTACTTCCCTGCTGCTGCTGCTGCTGCT 840  
DB 781 GTGATGACCTATTTAGTATFAGTACGATAGTACCTACTTCCCTGCTGCTGCTGCTGCTGCT 840  
QY 841 CTGCTGCTCATCTCAAGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 841 CTGCTGCTCATCTCAAGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 901 GCCTGTTTCCCGAGCCACAGGCGCTGCTACGACCCCGACAGATGTGCCCAAGCACTGCTC 960  
DB 901 GCCTGTTTCCCGAGCCACAGGCGCTGCTACGACCCCGACAGATGTGCCCAAGCACTGCTC 960  
QY 961 ACCATCCCGACCCACCCAGCAAGCAGGCGTGGAGTCCAGGCCACTTTCAGCTGCTC 1020  
DB 961 ACCATCCCGACCCACCCAGCAAGCAGGCGTGGAGTCCAGGCCACTTTCAGCTGCTC 1020  
QY 1021 CTTTCTCCGTCGATCGTCT 1080  
DB 1021 CTTTCTCCGTCGATCGTCT 1080  
QY 1081 TGACCCCTCCCTCCGCTGCTGCTTTCGATCAAAAGCTCTCAAAACCCCGCTGCTGCTGCT 1140  
DB 1081 TGACCCCTCCCTCCGCTGCTGCTTTCGATCAAAAGCTCTCAAAACCCCGCTGCTGCTGCT 1140  
QY 1141 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1141 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 GCTGCTGGGAGGCTTCCACCTTGGATCCAGAGGACCCCTCCACCTGCTGCTGCTGCTGCT 1260  
DB 1201 GCTGCTGGGAGGCTTCCACCTTGGATCCAGAGGACCCCTCCACCTGCTGCTGCTGCTGCT 1260  
QY 1261 GGCAGAAAGGCTACAGGCTGCTTCTCTCTCTCCCGACCCCGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 GGCAGAAAGGCTACAGGCTGCTTCTCTCTCTCCCGACCCCGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
DB 1321 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
QY 1381 CCGGGAGTCTCGAGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
DB 1381 CCGGGAGTCTCGAGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
QY 1441 CCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1441 CCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
QY 1501 AGAAGGAAACAAATAGGGGCTGGCGGATATAAAGTCACTTCTTCTTCTTCTTCTTCTTCT 1560  
DB 1501 AGAAGGAAACAAATAGGGGCTGGCGGATATAAAGTCACTTCTTCTTCTTCTTCTTCTTCT 1560  
QY 1561 TGACAGATGAACAAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGAGGACCAAGGC 1620  
DB 1561 TGACAGATGAACAAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGAGGACCAAGGC 1620  
QY 1621 CTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGCCAACCAT 1680



Db 1621 |||||CTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCAGTGACACTGTGACGGCCAACCAT 1680  
Qy 1681 GCACCTCCACAAAGGGAGACTTTTGGAAATGAAGACTAGCTCCTATGTATGATCAGGTAAAGA 1740  
Db 1681 GCACCTCCACAAAGGGAGACTTTTGGAAATGAAGACTAGCTCCTATGTATGATCAGGTAAAGA 1740  
Qy 1741 GCAAGGAGAGCTGCCAGGAGCAGCAGCTTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Db 1741 GCAAGGAGAGCTGCCAGGAGCAGCAGCTTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Qy 1801 GGGGCTCTAGGCCCCCATCTTCATTTCTTAGTTAAGAGAGCATTTTCCTCAGACTGCCA 1860  
Db 1801 GGGGCTCTAGGCCCCCATCTTCATTTCTTAGTTAAGAGAGCATTTTCCTCAGACTGCCA 1860  
Qy 1861 GGGGAGAGACTAGCCTAGCCTTCAGCAACCAAGGTTCTCTGGGAGCCCAAAGTTTATGG 1920  
Db 1861 GGGGAGAGACTAGCCTAGCCTTCAGCAACCAAGGTTCTCTGGGAGCCCAAAGTTTATGG 1920  
Qy 1921 GAGAAGGGCAAGACTTTTCATGGGAGAGAGAGGAGGCCCTGGGTAGAAACGCTTGGTG 1980  
Db 1921 GAGAAGGGCAAGACTTTTCATGGGAGAGAGAGGAGGCCCTGGGTAGAAACGCTTGGTG 1980  
Qy 1981 CTGTTCTTTGGGCTTTAAGACAAAGCGCTCATCTTGGCCCTCTACCTCTCTGATAGGCTT 2040  
Db 1981 CTGTTCTTTGGGCTTTAAGACAAAGCGCTCATCTTGGCCCTCTACCTCTCTGATAGGCTT 2040  
Qy 2041 GAGGTTTGGCCAAACCACTGTGGCTACAGGTGAGGGAAGAGACATCTTCTCCAGAG 2100  
Db 2041 GAGGTTTGGCCAAACCACTGTGGCTACAGGTGAGGGAAGAGACATCTTCTCCAGAG 2100  
Qy 2101 TGCATATGTTTCAAGGAAGTTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCC 2160  
Db 2101 TGCATATGTTTCAAGGAAGTTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCC 2160  
Qy 2161 TTTTAAAGCGGAAACAAGTAATTTTACAGTTTCTACTGGGGTTCTTCCACCCAGG 2220  
Db 2161 TTTTAAAGCGGAAACAAGTAATTTTACAGTTTCTACTGGGGTTCTTCCACCCAGG 2220  
Qy 2221 TGGGCGAGGCTTAGGAAGAGGTCATCTTAAAGCCACACATTTAGCTGCACCTGGCGTG 2280  
Db 2221 TGGGCGAGGCTTAGGAAGAGGTCATCTTAAAGCCACACATTTAGCTGCACCTGGCGTG 2280  
Qy 2281 CAGCCAAAAACAAGAACTGGGTGTTAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Db 2281 CAGCCAAAAACAAGAACTGGGTGTTAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Qy 2341 CTCATATGTGAAGGATAGAACTCCTCCTTACTCTTCTCCAAAAAGAGTGGGAAAGA 2400  
Db 2341 CTCATATGTGAAGGATAGAACTCCTCCTTACTCTTCTCCAAAAAGAGTGGGAAAGA 2400  
Qy 2401 ACCATCAAACTTTCTCCTCCTCCTACTTACCAACCCAGGAAAAACAGCAGGAGGGTGCTCA 2460  
Db 2401 ACCATCAAACTTTCTCCTCCTCCTACTTACCAACCCAGGAAAAACAGCAGGAGGGTGCTCA 2460  
Qy 2461 GGACTTAGGGACAGGTATAGCTTAGATGTGGTGAAGAACCAAGGAGAGCAGGAAGTTGTAA 2520  
Db 2461 GGACTTAGGGACAGGTATAGCTTAGATGTGGTGAAGAACCAAGGAGAGCAGGAAGTTGTAA 2520  
Qy 2521 ATCACTGGCTAATGAGAAAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580  
Db 2521 ATCACTGGCTAATGAGAAAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580  
Qy 2581 AGTTGCTCTCTGAAGTGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640  
Db 2581 AGTTGCTCTCTGAAGTGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640  
Qy 2641 GCTAGGGAAGTAGGTGATGCCAGGCCCTCAAGCTCTGTCTTCAGCCAGGAGCTTGAAGATT 2700  
Db 2641 GCTAGGGAAGTAGGTGATGCCAGGCCCTCAAGCTCTGTCTTCAGCCAGGAGCTTGAAGATT 2700  
Qy 2701 ATATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCCCTGGAAGATCAGGCA 2760  
|||||

Db 2701 ATATTGGGCACTGGCTCAATCTGTGTGACCAAGTATTTTCCCTTCCCTGGAAGATCAGGCA 2760  
Qy 2761 GGGTGCCATTCAATTGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTTGTAC 2820  
Db 2761 GGGTGCCATTCAATTGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTTGTAC 2820  
Qy 2821 TGTACCTTAGGGTCTTGGAAAGGAAAAACATGGAAATCAGGAATCTATAGACTGATAGGCC 2880  
Db 2821 TGTACCTTAGGGTCTTGGAAAGGAAAAACATGGAAATCAGGAATCTATAGACTGATAGGCC 2880  
Qy 2881 CTATCCACAAAGGGCCATGACTGGGAAAGGATATGGGAGCAAGAGAGAAATTTGGGATTTTA 2940  
Db 2881 CTATCCACAAAGGGCCATGACTGGGAAAGGATATGGGAGCAAGAGAGAAATTTGGGATTTTA 2940  
Qy 2941 GGGTGCACTACGCTCAACCCCTAAACTTTTGTGGCCTGGGCAATGTCTTGAGGCCCAGAC 3000  
Db 2941 GGGTGCACTACGCTCAACCCCTAAACTTTTGTGGCCTGGGCAATGTCTTGAGGCCCAGAC 3000  
Qy 3001 TGTAAAGCAGGCTCTGTGGCCTGTTTACCTGTCACCACTCTGCACCTGCTGCTTTGAG 3060  
Db 3001 TGTAAAGCAGGCTCTGTGGCCTGTTTACCTGTCACCACTCTGCACCTGCTGCTTTGAG 3060  
Qy 3061 ACTCCATCCAGCCCCAGGCAAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGAGG 3120  
Db 3061 ACTCCATCCAGCCCCAGGCAAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGAGG 3120  
Qy 3121 GTGAATCTGTGTACTGTGTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 3180  
Db 3121 GTGAATCTGTGTACTGTGTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 3180  
Qy 3181 TTAACACACAGATGTTTACAAAATAAAGATTATTTCAAACCAACCAAAAAA 3240  
Db 3181 TTAACACACAGATGTTTACAAAATAAAGATTATTTCAAACCAACCAAAAAA 3240  
Qy 3241 AAAAAAAAAAAAAAAAAA 3257  
Db 3241 AAAAAAAAAAAAAAAAAA 3257

RESULT 2  
AAK82285/c  
ID AAK82285 standard; DNA; 3240 BP.  
XX  
AC AAK82285;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37097.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209457P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.



CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 3240 BP; 713 A; 882 C; 887 G; 758 T; 0 U; 0 Other;

Query Match 82.8%; Score 2696.6; DB 4; Length 3240;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 271; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 473 TGGCATGGACTTCTATGAACAAGTGGCTGCACCAACCTCAACATTAAAGAGTCAATTCAC 532  
Db 3100 TGCCAGGGAGAGGAGGAGACACTGGACTAACCTGTGCCCTTTGGTTTCCAGTCAATTCAC 3041

Qy 533 GCGTCTGACAGAGTGTGTCTGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGGATGCG 592  
Db 3040 GCGTCTGACAGAGTGTGTCTGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGGATGCG 2981

Qy 593 TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGCAAAACCCGAGGGCCC 652  
Db 2980 TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGCAAAACCCGAGGGCCC 2921

Qy 653 AGCGAACTCTTTCGAAACCTGCTGGTGTGAGTCTCTGTGTGGGCGACCCCAACAGACACC 712  
Db 2920 AGCGAACTCTTTCGAAACCTGCTGGTGTGAGTCTCTGTGTGGGCGACCCCAACAGACACC 2861

Qy 713 CCTCTTCCCTCAGAGGCCGTGGGAGACAGAGGGAGCCGGGGCTTTGGCCCTGTGTGTGT 772  
Db 2860 CCTCTTCCCTCAGAGGCCGTGGGAGACAGAGGGAGCCGGGGCTTTGGCCCTGTGTGTGT 2801

Qy 773 CCTCTGCTGTATGACCTATTGAGTATCAGTAGCCACTACCCCTGCTGGCCCTGA 832  
Db 2800 CCTCTGCTGTATGACCTATTGAGTATCAGTAGCCACTACCCCTGCTGGCCCTGA 2741

Qy 833 GAGGGGCTCTGCTGTCTATCTCAAGCAGCCCTGTCCCGAGCCCTCCACCCCTGGAGTGGT 892  
Db 2740 GAGGGGCTCTGCTGTCTATCTCAAGCAGCCCTGTCCCGAGCCCTCCACCCCTGGAGTGGT 2681

Qy 893 CTTCTTCAGCCGTGTTTCCCGAGCCAGCGCTGTACAGACCCCAAGATGTCGCGAAGC 952  
Db 2680 CTTCTTCAGCCGTGTTTCCCGAGCCAGCGCTGTACAGACCCCAAGATGTCGCGAAGC 2621

Qy 953 ACTGTCTACCATCCGACCCACAGACACAGCCAGGCTGAGTCCAGGCCACTTTC 1012  
Db 2620 ACTGTCTACCATCCGACCCACAGACACAGCCAGGCTGAGTCCAGGCCACTTTC 2561

Qy 1013 AGCTGCTCTCTTCTCCGTGCAATCGTGTCTCTTCTCTGCTTTTCTCTCTCTCTCTCTCTCT 1072  
Db 2560 AGCTGCTCTCTTCTCCGTGCAATCGTGTCTCTTCTCTGCTTTTCTCTCTCTCTCTCTCT 2501

Qy 1073 TCTTTCTCTGACCCCTCCCTCCCGTGGTTCGTATCAAAAGCTCTCAAAACCCCGTCCC 1132  
Db 2500 TCTTTCTCTGACCCCTCCCTCCCGTGGTTCGTATCAAAAGCTCTCAAAACCCCGTCCC 2441

Qy 1133 CCGTGTCTCTGCTGTGTGACGTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1192  
Db 2440 CCGTGTCTCTGCTGTGTGACGTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2381

Qy 1193 GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCT 1252  
Db 2380 GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCT 2321

Qy 1253 CGTGGGTGGGCAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1312  
Db 2320 CGTGGGTGGGCAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2261

Qy 1313 ATGTGCCATGGGCTGCTCTCCCAAGTGCAGCTGCGAAAGTGAGGATCGAGGTAGAGGGA 1372  
Db 2260 ATGTGCCATGGGCTGCTCTCCCAAGTGCAGCTGCGAAAGTGAGGATCGAGGTAGAGGGA 2201

Qy 1373 AACAGCAACCGGGAGTGCTCTCGAGCCTGGGGCTGCCCTACTCTTACCCATTTCCCGACCA 1432  
Db 2200 AACGGCAACCGGGAGTGCTCTCGAGCCTGGGGCTGCCCTACTCTTACCCATTTCCCGACCA 2141

Qy 1433 GAGCTTTGGCCCTTGTCTTGGCTGCGCCGCTGCTCTTTTGGGGAAGTGAAGTCTCAGAGGAGG 1492  
Db 2140 GAGCTTTGGCCCTTGTCTTGGCTGCGCCGCTGCTCTTTTGGGGAAGTGAAGTCTCAGAGGAGG 2081

Qy 1493 TGCTTTCAGAGAAAGGAAACAAAATGAGGGGTGGCAGGAGTAAAAAGTCACTCTTCTCTCTCT 1552  
Db 2080 TGCTTTCAGAGAAAGGAAACAAAATGAGGGGTGGCAGGAGTAAAAAGTCACTCTTCTCTCT 2021

Qy 1553 ACTTCCCATGAGCATGAACACAAATTTCTCTCCACCTGGTCTCCAAATTTAAAGATGAGG 1612  
Db 2020 ACTTCCCATGAGCATGAACACAAATTTCTCTCCACCTGGTCTCCAAATTTAAAGATGAGG 1961

Qy 1613 ACCAAGCCCTGTGGGTACTCCAGGGGCAAGGAGCCCTGGGTGAGTGCACACACCTGTCTCAGG 1672  
Db 1960 ACCAAGCCCTGTGGGTACTCCAGGGGCAAGGAGCCCTGGGTGAGTGCACACACCTGTCTCAGG 1901

Qy 1673 CCAACCATGCACTCCACAAAGGGGAGCATTTTGGAAATGAAGGACTAGTCTCTATGTATCA 1732  
Db 1900 CCAACCATGCACTCCACAAAGGGGAGCATTTTGGAAATGAAGGACTAGTCTCTATGTATCA 1841

Qy 1733 GGTTAAGAGCAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGCAGAGGGAATGTAG 1792  
Db 1840 GGTTAAGAGCAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGCAGAGGGAATGTAG 1781

Qy 1793 CAACAGCAGGGCCCTCTTAGGCCCATCTTCCATTTCTTAGGTAAAGAGAGCATTTTCTCTCA 1852  
Db 1780 CAACAGCAGGGCCCTCTTAGGCCCATCTTCCATTTCTTAGGTAAAGAGAGCATTTTCTCTCA 1721

Qy 1853 GACTTCCAGCGGAGGAGTGAAGCTAGCTTTCAGCAACCAAGGTTCTCTGGGACCCAAA 1912  
Db 1720 GACTTCCAGCGGAGGAGTGAAGCTAGCTTTCAGCAACCAAGGTTCTCTGGGACCCAAA 1661

Qy 1913 GTTTATGGGAGAAAGGCAAGACTTTCATGGGAAGAGAAAGAGGCTCTGGGTAGAAAC 1972  
Db 1660 GTTTATGGGAGAAAGGCAAGACTTTCATGGGAAGAGAAAGAGGCTCTGGGTAGAAAC 1601

Qy 1973 GCTTGTGTCTGTCTCTTGTGGCTTTTAAGACAAAGCGCTCATCTTGGCTCTACCTCTCTG 2032  
Db 1600 GCTTGTGTCTGTCTCTTGTGGCTTTTAAGACAAAGCGCTCATCTTGGCTCTACCTCTCTG 1541

Qy 2033 ATAGGCTTGGGGTTTCCCAACACACTGTGGCTTACAGGTGGAGGGAAGAGAGTCTCTTC 2092  
Db 1540 ATAGGCTTGGGGTTTCCCAACACACTGTGGCTTACAGGTGGAGGGAAGAGAGTCTCTTC 1481

Qy 2093 CTCAGAGTGTATGTTTTCAGGAAGTCTTTTAAACCCCATATGGCCCAAGAGTAGTCTGTA 2152  
Db 1480 CTCAGAGTGTATGTTTTCAGGAAGTCTTTTAAACCCCATATGGCCCAAGAGTAGTCTGTA 1421

Qy 2153 GGAGGCCCTTTAAAGACGGAACAAAGTAAATTTACAGTTCTACTGGGGTTCTTGGCCACCG 2212  
Db 1420 GGAGGCCCTTTAAAGACGGAACAAAGTAAATTTACAGTTCTACTGGGGTTCTTGGCCACCG 1361

Qy 2213 TCCCAAGGTGGGAGGCTTAGGAAGAGGGTCAATTTTAAAGCCACACATTAAGTGCACCTG 2272  
Db 1360 TCCCAAGGTGGGAGGCTTAGGAAGAGGGTCAATTTTAAAGCCACACATTAAGTGCACCTG 1301

Qy 2273 CGTGGCTGACGCAAAAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAAT 2332  
Db 1300 CGTGGCTGACGCAAAAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAAT 1241

Qy 2333 CCAGGGCACTCATATGTGAAGGATAAGAACTCACTTCTCTTCTTCTCTCTCTCTCTCTCTCTCT 2392  
Db 1240 CCAGGGCACTCATATGTGAAGGATAAGAACTCACTTCTCTTCTCTCTCTCTCTCTCTCTCTCT 1181

Qy 2393 GGGAAAGAACCATCAAAACCTTCTCTGACTTACCAACACAGGAAACACAGAGGAGG 2452  
Db 1180 GGGAAAGAACCATCAAAACCTTCTCTGACTTACCAACACAGGAAACACAGAGGAGG 1121

Qy 2453 GTGGCTCAGGACTTATGGGACAGGGGTATAGCTTGTGTGAAAGAAAGGAGAGGAGG 2512

Db 1120 GTGGCTCAGGACTTAGGCAGCAGGCTATAGCTTAGATGGTGAAGCAAGCAGCAGGA 1061  
Qy 2513 AGTTGTAATCACTGGCTTAATGAGAAAGAGACAGCTAACTCTAGGATGAAGCTGTGAC 2572  
Db 1060 AGTTGTAATCACTGGCTTAATGAGAAAGAGACAGCTAACTCTAGGATGAAGCTGTGAC 1001  
Qy 2573 TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632  
Db 1000 TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 941  
Qy 2633 ACATGGGCTTAGGAAAGTAGGATGCCAGCCCTCAAGTCTGTCTTCAGCAGGACTT 2692  
Db 940 ACATGGGCTTAGGAAAGTAGGATGCCAGCCCTCAAGTCTGTCTTCAGCAGGACTT 881  
Qy 2693 GAGNAGTATATTGGGAGTGGCTCCATCTGTGGACAGTATTTACGCTTCCCTGAAG 2752  
Db 880 GAGNAGTATATTGGGAGTGGCTCCATCTGTGGACAGTATTTACGCTTCCCTGAAG 821  
Qy 2753 ATCAGGAGGCTGCATTCTTCTTCTCTCTAGCCCTCAGGAAAGAGGACTAT 2812  
Db 820 ATCAGGAGGCTGCATTCTTCTCTCTCTAGCCCTCAGGAAAGAGGACTAT 761  
Qy 2813 ATTGTACTGTACCTTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTTATAGACT 2872  
Db 760 ATTGTACTGTACCTTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTTATAGACT 701  
Qy 2873 GATAGGCCCTATCCACAGGCCCATGCTGGGAAAGGATATGGAGCAGAGGAAATTG 2932  
Db 700 GATAGGCCCTATCCACAGGCCCATGCTGGGAAAGGATATGGAGCAGAGGAAATTG 641  
Qy 2933 GGATTTTAGGCTGAGCTAGCTCACCCTAACTTTTGGTGGCTGGGGCATGCTTTGAG 2992  
Db 640 GGATTTTAGGCTGAGCTAGCTCACCCTAACTTTTGGTGGCTGGGGCATGCTTTGAG 581  
Qy 2993 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTTTTACTGTCACCACTCTGCACCTGCT 3052  
Db 580 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTTTTACTGTCACCACTCTGCACCTGCT 521  
Qy 3053 GTCTTGAGACTCCATCCAGCCAGCCAGCCAGCTCTCTGAGCCTCCATCTCTCCC 3112  
Db 520 GTCTTGAGACTCCATCCAGCCAGCCAGCCAGCTCTCTGAGCCTCCATCTCTCCC 461  
Qy 3113 TGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCCATATATGAATTTGAGCAGGGTT 3172  
Db 460 TGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCCATATATGAATTTGAGCAGGGTT 401  
Qy 3173 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTTTCAAAACCACC 3223  
Db 400 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTTTCAAAACCACC 350

RESULT 3

AAD47169  
ID AAD47169 standard; DNA; 28770 BP.

AC AAD47169;

XX 24-FEB-2003 (first entry)

DE Human Ras-like protein encoding gene.

XX Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
XX immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
XX Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
XX osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
XX irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
XX pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
XX Sjogren's syndrome; infection; transgenic; gene therapy; neotropic; gout;  
XX neuroprotective; vasotropic; cytotostatic; dermatological; nephrotropic;  
XX ophthalmological; tranquilliser; cancer; stroke; Grave's disease; AIDS;  
XX asthma; anaemia; drug screening; gene; chromosome 14; ds.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH variation replace(206, -)  
FT /tag= a  
FT /standard\_name= "Single nucleotide polymorphism"  
FT CDS 2044..25772  
FT /tag= b  
FT /product= "Human Ras-like protein"  
FT exon 2044..2167  
FT /tag= c  
FT intron 2168..21554  
FT /tag= d  
FT variation replace(4963,T)  
FT /tag= e  
FT variation replace(8175,A)  
FT /tag= f  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(10515,C)  
FT /tag= g  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(13034,C)  
FT /tag= h  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(13781,C)  
FT /tag= i  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(14050,C)  
FT /tag= j  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(14273, -)  
FT /tag= k  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(17582,C)  
FT /tag= l  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(17700,T)  
FT /tag= m  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(18074, TC)  
FT /tag= n  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(19328,T)  
FT /tag= o  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(19570,G)  
FT /tag= p  
FT /standard\_name= "Single nucleotide polymorphism"  
FT variation replace(20892,T)  
FT /tag= q  
FT /standard\_name= "Single nucleotide polymorphism"  
FT exon 21555..21615  
FT /tag= r  
FT intron 21616..22462  
FT /tag= s  
FT exon 22463..22523  
FT /tag= t  
FT intron 22524..22974  
FT /tag= u  
FT exon 22975..23052  
FT /tag= v  
FT intron 23053..23711  
FT /tag= w  
FT exon 23712..23801  
FT /tag= x  
FT intron 23802..25392  
FT /tag= y  
FT exon 25393..25458  
FT /tag= z  
FT intron 25459..25613  
FT /tag= aa  
FT exon 25614..25769  
FT /tag= ab



1613 ACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGGTCACTGACACTGTGACG 1672  
Db |  
26702 ACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGGTCACTGACACTGTGACG 26761  
Qy |  
1673 CCAACATGACCTCCAAAGAGGGAGCATTTGGAATGAAGGACTAGCTCCTATGATCA 1732  
Db |  
26762 CCAACATGACCTCCAAAGAGGGAGCATTTGGAATGAAGGACTAGCTCCTATGATCA 26821  
Qy |  
1733 GGTATAGAGCAAGGAGAGCTGGCCAGGAGCAGCAGTTTGCACAGCAGAGGGGAATGTAG 1792  
Db |  
26822 GGTATAGAGCAAGGAGAGCTGGCCAGGAGCAGCAGTTTGCACAGCAGAGGGGAATGTAG 26881  
Qy |  
1793 CAACAGCAGGCGCTCCTPAGGCCCATCTTCCATTTCTTAGGTAAAGAGACATTTCTCTCA 1852  
Db |  
26882 CAACAGCAGGCGCTCCTPAGGCCCATCTTCCATTTCTTAGGTAAAGAGACATTTCTCTCA 26941  
Qy |  
1853 GACTCCAGGCGGAGGACTGAGCCTAGCCTTCCAGCAACCAAGGTTCTCTGGGACCCAAA 1912  
Db |  
26942 GACTCCAGGCGGAGGACTGAGCCTAGCCTTCCAGCAACCAAGGTTCTCTGGGACCCAAA 27001  
Qy |  
1913 GTTATGGAGAGGGCAAGACTTTCATGGAGAGAGAGAGGCCCTGGGTAGAAC 1972  
Db |  
27002 GTTATGGAGAGGGCAAGACTTTCATGGAGAGAGAGAGGCCCTGGGTAGAAC 27061  
Qy |  
1973 GCTTGGTGCTGTTCTCTTGGCCCTTAAAGCAAGCGCTCATCTTGGCCCTTACCTCCTG 2032  
Db |  
27062 GCTTGGTGCTGTTCTCTTGGCCCTTAAAGCAAGCGCTCATCTTGGCCCTTACCTCCTG 27121  
Qy |  
2033 ATAGGCTTGAAGGTTTCCCAACCACTGTGGCTACAGGTGGAGGGAAGAGGACTCCTTC 2092  
Db |  
27122 ATAGGCTTGAAGGTTTCCCAACCACTGTGGCTACAGGTGGAGGGAAGAGGACTCCTTC 27181  
Qy |  
2093 CTCAGAGTCTATGTTACAGAGTTTCTTAAACCCATATGCCCCAAGAGTAGCTCGTA 2152  
Db |  
27182 CTCAGAGTCTATGTTACAGAGTTTCTTAAACCCATATGCCCCAAGAGTAGCTCGTA 27241  
Qy |  
2153 GGAGGCCCTTTAAGAGCGAAGTAATTTACCAGTTCTACTGGGGTCTCTGCCACCG 2212  
Db |  
27242 GGAGGCCCTTTAAGAGCGAAGTAATTTACCAGTTCTACTGGGGTCTCTGCCACCG 27301  
Qy |  
2213 TCCCAAGTGGCGAGGCTAGGAAGGGTCAATTTCTTAAGCCACACATAGCTGCACTG 2272  
Db |  
27302 TCCCAAGTGGCGAGGCTAGGAAGGGTCAATTTCTTAAGCCACACATAGCTGCACTG 27361  
Qy |  
2273 CGTGGCTGACGACCAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 2332  
Db |  
27362 CGTGGCTGACGACCAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 27421  
Qy |  
2333 CCAGGGCACTCATATGTGAAGGATAAGAACTCACTTCTTCTCTTCCCAAAAGAAAGTG 2392  
Db |  
27422 CCAGGGCACTCATATGTGAAGGATAAGAACTCACTTCTTCTCTTCCCAAAAGAAAGTG 27481  
Qy |  
2393 GGGAAAGAACATCAAACTTCTCTGACTTACCAAAACAGGAAACAGCAGGAGG 2452  
Db |  
27482 GGGAAAGAACATCAAACTTCTCTGACTTACCAAAACAGGAAACAGCAGGAGG 27541  
Qy |  
2453 GTGGCTCAGGACTTAGGCAAGGATAGCTTAGATGGTGGAAAGCAAGGAGCAGGA 2512  
Db |  
27542 GTGGCTCAGGACTTAGGCAAGGATAGCTTAGATGGTGGAAAGCAAGGAGCAGGA 27601  
Qy |  
2513 AGTTGTAATCACTGGCTAATGAGAAAGGAGACAGCTTAATCTAGGATGAAGCTGTGAC 2572  
Db |  
27602 AGTTGTAATCACTGGCTAATGAGAAAGGAGACAGCTTAATCTAGGATGAAGCTGTGAC 27661  
Qy |  
2573 TAGCTCGAGTGTCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCATC 2632  
Db |  
27662 TAGCTCGAGTGTCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCATC 27721  
Qy |  
2633 ACCTGGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTT 2692  
Db |  
27722 ACCTGGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTT 27781

2693 GAGAAGTTATATGGCAGTGGCTCCAATCTGTGGACCAAGTATTTAGCTTTCCCTGAAG 2752  
Db |  
27782 GAGAAGTTATATGGCAGTGGCTCCAATCTGTGGACCAAGTATTTAGCTTTCCCTGAAG 27841  
Qy |  
2753 ATCAGCAGGAGTGGCAATTCATTTCTCTCTAGCCCCCTCAGAAAGAGGACTAT 2812  
Db |  
27842 ATCAGCAGGAGTGGCAATTCATTTCTCTCTAGCCCCCTCAGAAAGAGGACTAT 27901  
Qy |  
2813 ATTTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAATCAGGATTTCTATAGACT 2872  
Db |  
27902 ATTTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAATCAGGATTTCTATAGACT 27961  
Qy |  
2873 GATAGCCCTTATCCAAAGGGCCATGACTGGGAAAAGGTATGGAGCAGAGGAGATTG 2932  
Db |  
27962 GATAGCCCTTATCCAAAGGGCCATGACTGGGAAAAGGTATGGAGCAGAGGAGATTG 28021  
Qy |  
2933 GATTTTAGGAGTGCAGTACGCTCACCTAAACTTTTGTGGCTGGGGCATGTCTTGAG 2992  
Db |  
28022 GATTTTAGGAGTGCAGTACGCTCACCTAAACTTTTGTGGCTGGGGCATGTCTTGAG 28081  
Qy |  
2993 GCCCAGACTGTTAAGCAGGCTCTGTGGCTGTCTTACTGTCTCACCACTCTGCACCTGCT 3052  
Db |  
28082 GCCCAGACTGTTAAGCAGGCTCTGTGGCTGTCTTACTGTCTCACCACTCTGCACCTGCT 28141  
Qy |  
3053 GTCTTGAGACTCCATCCAGCCCGAGGACGCCACCTGCTCTGAGCCTCCACTATCTCCC 3112  
Db |  
28142 GTCTTGAGACTCCATCCAGCCCGAGGACGCCACCTGCTCTGAGCCTCCACTATCTCCC 28201  
Qy |  
3113 TGTGACGGGTGAATCTGCTGCTACTGTCTCGGGTCCATATATGATTTGAGCAGGGTT 3172  
Db |  
28202 TGTGACGGGTGAATCTGCTGCTACTGTCTCGGGTCCATATATGATTTGAGCAGGGTT 28261  
Qy |  
3173 CATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCAACC 3223  
Db |  
28262 CATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCAACC 28312  
RESULT 4  
AAC75813  
ID AAC75813 standard; cDNA; 1666 BP.  
XX  
AAC75813;  
XX  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORF1368 polynucleotide sequence SEQ ID NO:2735.  
XX  
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnery; antipsoriatric; antiparkinsonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive; ss.  
OS Homo sapiens.  
XX  
XX W020058473-A2.  
XX  
PD 05-OCT-2000.  
XX  
PF 31-MAR-2000; 2000WO-US008621.  
XX  
PR 31-MAR-1999; 99US-0127607P.  
PR 02-APR-1999; 99US-0127636P.  
PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.  
XX (CURA-) CURAGEN CORP.  
XX Shimkets RA, Leach M;  
XX WPI; 2000-602362/57.  
DR P-PSDB; AAB41604.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 5; Page 1978-1979; 5507pp; English.  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORF open reading frames 1 to 3161. The ORF  
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
CC antiproliferative; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORF-associated disorder. The nucleic acids can be used to express ORF  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 1666 BP; 333 A; 529 C; 461 G; 341 T; 0 U; 2 Other;

Query Match 48.1%; Score 1568.2; DB 3; Length 1666;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1598; Conservative 0; Mismatches 3; Indels 18; Gaps 1;

Qy 2 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCGCTGGCCCGCTGATGCGGAGCAGTACGA 61  
Db 48 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCGCTGGCCCGCTGATGCGGAGCAGTACGA 107  
Qy 62 TGTGCTGTTCGGCTGTCTGTGATGCGGGAATCCGGGGTGGGGAAGACCTGCGCTGTGTG 121  
Db 108 TGTGCTGTTCGGCTGTCTGTGATGCGGGAATCCGGGGTGGGGAAGACCTGCGCTGTGTG 167  
Qy 122 CCGCTTCCACGACACGAGTTCCACTCTCCGACATCTCCACCATCGGTGTGACTTTAA 181  
Db 168 CCGCTTCCACGACACGAGTTCCACTCTCCGACATCTCCACCATCGGTGTGACTTTAA 227  
Qy 182 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACATGCGAGG 241  
Db 228 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACATGCGAGG 287  
Qy 242 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCCGAGGGGATATTTT 301  
Db 288 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCCGAGGGGATATTTT 347  
Qy 302 GGTCTATGACATTAGCAGCGCGCTTACCAGCATCATGATGAGTGGGTCACTGACGT 361  
Db 348 GGTCTATGACATTAGCAGCGCGCTTACCAGCATCATGATGAGTGGGTCACTGACGT 407  
Qy 362 GGATGAGTACACACAGAGCGCTCCAGAGATCTTATGGGAATAGGCTGATGAGGA 421  
Db 408 GGATGAGTACACACAGAGCGCTCCAGAGATCTTATGGGAATAGGCTGATGAGGA 467  
Qy 422 GCAGAAACGGCAGGTGGGGAAGAGAGCAGGGCAG-----CAGCTGGC 463  
Db 468 GCAGAAACGGCAGGTGGGGAAGAGAGCAGGGCAG-----CAGCTGGC 527

Qy 464 GAAGAGCTATGGCATGACCTTCTATGAAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA 523  
Db 528 GAAGAGCTATGGCATGACCTTCTATGAAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA 587  
Qy 524 GTCAATCAACGCTGTGACAGAGCTGTGTGAGGCCCCATAGGAAGAGAGCTGGAAGGCT 583  
Db 588 GTCAATCAACGCTGTGACAGAGCTGTGTGAGGCCCCATAGGAAGAGAGCTGGAAGGCT 647  
Qy 584 CCGATCCGTCGACAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGAGGAGGAGGAGG 643  
Db 648 CCGATCCGTCGACAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGAGGAGGAGGAGG 707  
Qy 644 CCGAGGCCCCAGCGAACTCTTCGAAAACCTGCTGTGTGTGAGTCTCTGTGTGGGACACCCCA 703  
Db 708 CCGAGGCCCCAGCGAACTCTTCGAAAACCTGCTGTGTGTGAGTCTCTGTGTGGGACACCCCA 767  
Qy 704 CACGACACCCCTCTTCCCTCAGGAGGCGCTGTGGGACAGACAGGAGGAGGAGGAGGAGGAGG 763  
Db 768 CACGACACCCCTCTTCCCTCAGGAGGCGCTGTGGGACAGACAGGAGGAGGAGGAGGAGGAGG 827  
Qy 764 TGCTGTGTCTCTCGTGTGATGACCTATTGAGTATCAGTAGCCACTACTCCCTCCCTGCC 823  
Db 828 TGCTGTGTCTCTCGTGTGATGACCTATTGAGTATCAGTAGCCACTACTCCCTCCCTGCC 887  
Qy 824 TGGCCCTGAGAGCGGCTCTGCTGTCATCTCAAGCAGGCCCCCTGTCCCGAGCCCCGTCACCC 883  
Db 888 TGGCCCTGAGAGCGGCTCTGCTGTCATCTCAAGCAGGCCCCCTGTCCCGAGCCCCGTCACCC 947  
Qy 884 TGGAGTGGTCTTCTTACAGCTGTTCCTCCAGCCACAGGCGCTGTACGACCCCCAGATGT 943  
Db 948 TGGAGTGGTCTTCTTACAGCTGTTCCTCCAGCCACAGGCGCTGTACGACCCCCAGATGT 1007  
Qy 944 GCAGCAGACCTGTCTCACCATCCGACCCAGACCCAGACAGGAGGAGGAGGAGGAGGAGGAGG 1003  
Db 1008 GCAGCAGACCTGTCTCACCATCCGACCCAGACCCAGACAGGAGGAGGAGGAGGAGGAGGAGG 1067  
Qy 1004 GCACCTTTACAGTGTCTCTTCTCCGTCATCGTGTCTTCTCTCTGCTTTTCTCTCTCTTC 1063  
Db 1068 GCACCTTTACAGTGTCTCTTCTCCGTCATCGTGTCTTCTCTCTGCTTTTCTCTCTCTTC 1127  
Qy 1064 CCCCACCTTCTTCTCTGACCCCTCCCTCCGTCGCTGTTCCTCTCTCTCTCTCTCTCTCTCTCT 1123  
Db 1128 CCCCACCTTCTTCTCTGACCCCTCCCTCCGTCGCTGTTCGTCATCAAAAGCTCTCTCTCTCT 1187  
Qy 1124 CCCCCTCCCGGTGTCTGT 1183  
Db 1188 CCCCCTCCCGGTGTCTGT 1247  
Qy 1184 CAAGGGATGGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCTCC 1243  
Db 1248 CAAGGGATGGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCTCC 1307  
Qy 1244 ACCCTGCTCTGTGGGTGGGCCAAAGGCTACAGGCTGTCTTCTCTCTCTCTCTCTCTCTCTCT 1303  
Db 1308 ACCCTGCTCTGTGGGTGGGCCAAAGGCTACAGGCTGTCTTCTCTCTCTCTCTCTCTCTCTCT 1367  
Qy 1304 CTGTCCCTCATGTGCGATGGGCTGCTCCCTCCAGTGTCTCGGAAAGTGGAGCATCGAGG 1363  
Db 1368 CTGTCCCTCATGTGCGATGGGCTGCTCCCTCCAGTGTCTCGGAAAGTGGAGCATCGAGG 1427  
Qy 1364 TAGGAGGGAACAGCAACCGGGAGTCTCGAGGCTGGGGCTGCCCTTACCTCTACCCATT 1423  
Db 1428 TAGGAGGGAACAGCAACCGGGAGTCTCGAGGCTGGGGCTGCCCTTACCTCTACCCATT 1487  
Qy 1424 CCCCAGCAGAGCTTTTGGCTTGTGGCTGCGCGCTGCTCTCTTTTGGGGAACTGAGCTC 1483  
Db 1488 CCCCAGCAGAGCTTTTGGCTTGTGGCTGCGCGCTGCTCTCTTTTGGGGAACTGAGCTC 1547  
Qy 1484 AGAGGCGAGGTCTTCAGAGAGGAAAACAAATAGAGGGGTGCGAGGATATAAAGTCACT 1543  
Db 1548 AGAGGCGAGGTCTTCAGAGAGGAAAACAAATAGAGGGGTGCGAGGATATAAAGTCACT 1607









[illegible]

Db	1501	TTTTTTTTTTTTTTAGTTTTTTTTTTAGAAAGGAGTATATATTGTATTGTATTTTTTAGGGTT	1560
Qy	2836	CTGGAAGGGAACATCGAATCAGGATTCATAGACTGATAGGCCCTATCCACAAGGCC	2895
Db	1561	TTGGAAGGGAATATATGGAATTAGGATTTATAGATTGATAGTTTTATTATTAAGGTT	1620
Qy	2896	ATGACTGGGAAAAAGGTATGGGACGAGAAGAGAAATTTGGGATTTTAGGGTGCACCTACGCT	2955
Db	1621	ATGATTTGGGAAAAAGGTATGGGAGTAGAAGAGAAATTTGGGATTTTAGGGTGTATTACGTT	1680
Qy	2956	CACCTAAACTTTTGGTGGCTCGGGCGATGCTCTGAGGCCACAGCTGTTAAAGCAGGCTCT	3015
Db	1681	TATTTTAAATTTTGGTGGTTTGGGTAAGTTTTGAGTTTAGATTGTTAAATTAGGTTT	1740
Qy	3016	GCTGGCTGTTTTACTTCGTCAACCACTCTGCACCTGCTGCTTGAGACTCCATCCAGCCCC	3075
Db	1741	GTGGTTTGTATTTCGTATATTATTTCGTATTTGTTGTTTCGAGATTTATTATTAGTTT	1800
Qy	3076	AGGCAGCCACCTGCTCTCGAGCCTCCATATCTCCCTGTGACGGGTGAATTCGTGTAC	3135
Db	1801	AGGTACGTTATTTGTTTTCGAGTTTTATTATTTTTTTGTGACGGGTGAATTTTCGTGTAT	1860
Qy	3136	TGTGTCCTGGGTCCATATATGAATGTGACAGGGTTCATCTATTTTAAACACAGATGTT	3195
Db	1861	TGTGTTCCGGGTTATATATGAATTCGTAGTAGGGTTATTATTTTAAATATAGATGTT	1920
Qy	3196	TACAAATAAGATTTATTTCAACCACC	3223
Db	1921	TATAAAATAAAGATTTATTTTAAATATC	1948

RESULT 7  
ABK40070/c  
ID ABK40070 standard: DNA: 7924 BP.

RESULT 7	
ABK40070/c	
ID	ABK40070 standard; DNA; 7924 BP.
XX	
XX	
AC	ABK40070;
AC	
XX	
XX	
DT	21-MAY-2002 (first entry)
DE	
DE	Human chemically pretreated gene sequence #76 strand 2



QY 3078 GCAGGACCTGCTCTGAGCTCCACTATCTCCCTGTGAGGGTGAACCTTCGTACTG 3137  
DB 6122 ACAGGACCTACTCTCTAAACCTCCACTATCTCCCTATACGAATAAATCTCTACTA 6063  
QY 3138 TGCTCGGGTCCATATATATGATGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTA 3197  
DB 6062 TATCTGAATCCAT 6003  
QY 3198 CAAATATAGATTTTCAACACCCAAACCAACCAACCAACCAACCAACCAACCA 3257  
DB 6002 CAAATATATATATTTTCAACACCCAGATATATATATATATATATATATATAT 5943

RESULT 8  
ABL34139/c  
ID ABL34139 standard; DNA; 7924 BP.  
XX  
AC ABL34139;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 2112.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianemic; cytostatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antithyroid; antidiabetic; antipeoriatic;  
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
KW ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 02-JUL-2001; 2001WO-EP007537.  
XX  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2002-130909/17.  
XX  
PT Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
PS Claim 1; SEQ ID NO 2112; 32pp + Sequence Listing; German.  
XX  
CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
SQ Sequence 7924 BP; 1973 A; 83 C; 1827 G; 4041 T; 0 U; 0 Other;  
Query Match 35.6%; Score 1160.8; DB 6; Length 7924;  
Best Local Similarity 74.1%; Pred No. 1.3e-271;  
Matches 1468; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 1278 TGCTTCTCTCTTCCCAACCCCACTGTCCTCATGTGCAATGGCTGCTCCCAAG 1337  
DB 7922 TACTTCTCTCTTCCCAACCCCACTGTCCTCATATATATATATATATATATATAT 7863

QY 1338 TGACCTGCGAAAGTGGAGCATCGAGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGC 1397  
DB 7862 TAACTACGAAATTAACATCGAATTAACAAACGCAACCAACCAACCAACCAACCA 7803  
QY 1398 CTGGGCTGCT 1457  
DB 7802 CTAAAACTACT 7743  
QY 1458 GCTGCT 1517  
DB 7742 GCTTCT 7683  
QY 1518 GGGGTGGCAGGATATAAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1577  
DB 7682 AATAATACAAATATAAATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7623  
QY 1578 TTCT 1637  
DB 7622 TTCT 7563  
QY 1638 GCAAGGAGCCCTGGGGTCACTGACACTGTGAGGCAACCACTGCACTCCACAAAG 1697  
DB 7562 ACAAATAAACCCTTAAATCAATCACTATCAACCAACCACTATCACTATCACT 7503  
QY 1698 GCATTTGGAAATGAAGGACTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1757  
DB 7502 ACATTTAAATATAAATACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7443  
QY 1758 AGGACAGCAGTTTGCACAGCAGAGGGGATGTAGCAACAGAGGGCTCTCTAGG 1817  
DB 7442 AAAAAACAATTTTACACAAACCAACCAACCAACCAACCAACCAACCAACCA 7383  
QY 1818 TCTTCT 1877  
DB 7382 TCTTCT 7323  
QY 1878 AGCTTTCAGCAACCAAGTCTCTCTGGGACCCAAAGTTTATGGGAGAGGGCAAG 1937  
DB 7322 ACCTTTCACACCAACCAATCTCTCTTAAACCCCAATTTTATAAATAAATAA 7263  
QY 1938 CATGGGAG 1997  
DB 7262 CATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7203  
QY 1998 TAAGACAAAGCGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2057  
DB 7202 TAAACCAAAACGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7143  
QY 2058 ACTGTGCTACAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2117  
DB 7142 ACTATACTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7083  
QY 2118 TCTTTAAACCCATATGGCCCAAGAGTAGCTCTGAGGAGGCGCTTTTAAAGCGA 2177  
DB 7082 TCTTTAAACCCATATATACCCCAAAATAAATCTCTGTAATAAATAAATAAATA 7023  
QY 2178 TAAATTTACAGTCT 2237  
DB 7022 TAAATTTACCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6963  
QY 2238 GAGGCTCATTTCTTAAAGCCACACATTTAGCTGCTGCTGCTGCTGCTGCTGCT 2297  
DB 6962 AAAAACTATTTTAAACCAACCACTTACTACTACTACTACTACTACTACTACT 6903  
QY 2298 TGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGCACTCATATGTGAGG 2357  
DB 6902 TAAATATTAATAATTTCTCACTTAAACCAAAATCCAAACCACTCATATATA 6843  
QY 2358 AGACCT 2417  
DB 6842 AAAAACT 6783



PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234957P.	XX	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234958P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	26-SEP-2000;	2000US-0235484P.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235834P.	XX	WPI; 2001-465460/50.	
PR	29-SEP-2000;	2000US-0235836P.	XX	P-PSDB; AAU17136.	
PR	29-SEP-2000;	2000US-0236327P.	DR	Novel polypeptides useful for diagnosing, treating, preventing and/or	
PR	29-SEP-2000;	2000US-0236367P.	DR	prognosing disorders related to the proteins, including cancers, immune	
PR	29-SEP-2000;	2000US-0236368P.	XX	disorders and neuronal disorders.	
PR	29-SEP-2000;	2000US-0236369P.	PS	Claim 1; SEQ ID NO 88; 880pp; English.	
PR	29-SEP-2000;	2000US-0236370P.	XX	The invention relates to novel isolated polypeptides (I), and	
PR	02-OCT-2000;	2000US-0236802P.	CC	polynucleotides (II). (I), (II) and the antibody to (I) are useful for	
PR	02-OCT-2000;	2000US-0237037P.	CC	diagnosing, preventing and treating diseases including immune system	
PR	02-OCT-2000;	2000US-0237038P.	CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune	
PR	02-OCT-2000;	2000US-0237039P.	CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ	
PR	02-OCT-2000;	2000US-0241785P.	CC	transplant rejections and graft versus host disease, infectious diseases	
PR	20-OCT-2000;	2000US-0241786P.	CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and	
PR	20-OCT-2000;	2000US-0241808P.	CC	other blood-related disorders (sickle cell anaemia), myeloproliferative	
PR	20-OCT-2000;	2000US-0241809P.	CC	disorders, primary haematopoietic disorders, hyperproliferative disorders	
PR	20-OCT-2000;	2000US-0241826P.	CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.	
PR	01-NOV-2000;	2000US-0244617P.	CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities	
PR	08-NOV-2000;	2000US-0246474P.	CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.	
PR	08-NOV-2000;	2000US-0246475P.	CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),	
PR	08-NOV-2000;	2000US-0246476P.	CC	respiratory disorders, dermatological disorders, in wound healing,	
PR	08-NOV-2000;	2000US-0246477P.	CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's	
PR	08-NOV-2000;	2000US-0246478P.	CC	disease), reproductive system disorders, gastrointestinal disorder	
PR	08-NOV-2000;	2000US-0246523P.	CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of	
PR	08-NOV-2000;	2000US-0246524P.	CC	B-cell responsiveness to pathogens, activators of T-cells, to induce	
PR	08-NOV-2000;	2000US-0246525P.	CC	higher affinity antibodies, and as a means to induce tumour proliferation	
PR	08-NOV-2000;	2000US-0246526P.	CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-	
PR	08-NOV-2000;	2000US-0246527P.	CC	AAS27850 represent novel signal transduction pathway protein coding	
PR	08-NOV-2000;	2000US-0246528P.	CC	sequences and PCR primers of the invention	
PR	08-NOV-2000;	2000US-0246532P.	XX		
PR	08-NOV-2000;	2000US-0246609P.		Query Match	35.6%; Score 1160.6; DB 4; Length 2021;
PR	08-NOV-2000;	2000US-0246610P.		Best Local Similarity	99.2%; Pred. No. 7.9e-272;
PR	08-NOV-2000;	2000US-0246611P.		Matches 1166; Conservative	0; Mismatches 9; Indels 0; Gaps 0;
PR	08-NOV-2000;	2000US-0246613P.			
PR	17-NOV-2000;	2000US-0249207P.	Qy	2	GCCGGCTGCCCGCCCGCAGTTCCCGGCCCGCCCGCCCGCCAGTCATGCCGAGCAGTACGA 61
PR	17-NOV-2000;	2000US-0249208P.	Db	38	GCCCGCTGCCCGCCCGCAGTTCCCGGCCCGCCCGCCCGCCAGTCATGCCGAGCAGTACGA 97
PR	17-NOV-2000;	2000US-0249210P.	Qy	62	TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCCTGTGTG 121
PR	17-NOV-2000;	2000US-0249211P.	Db	98	TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCCTGTGTG 157
PR	17-NOV-2000;	2000US-0249212P.	Qy	122	CCGCTTACCGCAACGAGTTCCATCTCTCGCACATCTCCACCATCGGTGTGACTTTAA 181
PR	17-NOV-2000;	2000US-0249215P.	Db	158	CCGCTTACCGCAACGAGTTCCATCTCTCGCACATCTCCACCATCGGTGTGACTTTAA 217
PR	17-NOV-2000;	2000US-0249217P.	Qy	182	GATGAGACCATAGAGTAGACCGGCATCAAGTCGGATACAGATCTGGGACACTGCAGG 241
PR	17-NOV-2000;	2000US-0249218P.	Db	218	GATGAGACCATAGAGTAGACCGGCATCAAGTCGGATACAGATCTGGGACACTGCAGG 277
PR	17-NOV-2000;	2000US-0249219P.	Qy	242	GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCCGAGGGGATATTTT 301
PR	17-NOV-2000;	2000US-0249224P.	Db	278	GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCCGAGGGGATATTTT 337
PR	01-DEC-2000;	2000US-0250160P.	Qy	302	GGTCTATGACATTAGACGAGCGCTTTACACGACATCATGAGTGGGTGAGTACGT 361
PR	05-DEC-2000;	2000US-0250391P.	Db	338	GGTCTATGACATTAGACGAGCGCTTTACACGACATCATGAGTGGGTGAGTACGT 397
PR	05-DEC-2000;	2000US-0251988P.	Qy	362	GGATGAGTACGACCGACGAGCGCTTTATTTGGGAATAGCGTATGATGAGGA 421
PR	05-DEC-2000;	2000US-0256719P.	Db	398	GGATGAGTACGACCGACGAGCGCTTTATTTGGGAATAGCGTATGATGAGGA 457
PR	06-DEC-2000;	2000US-0251479P.			
PR	08-DEC-2000;	2000US-0251856P.			
PR	08-DEC-2000;	2000US-0251868P.			
PR	08-DEC-2000;	2000US-0251869P.			
PR	08-DEC-2000;	2000US-0251989P.			











CC different libraries may represent different parts of the same unknown  
CC gene, distorting the estimated frequency of occurrence in a particular  
CC tissue. AA243260-243309 represent expressed sequence tag (EST) fragments  
CC isolated from a human bladder tumour cDNA library which encode the  
CC proteins represented in AA243260-243309  
XX  
SQ Sequence 895 BP; 249 A; 205 C; 228 G; 213 T; 0 U; 0 Other;

Query Match 26.6%; Score 866.2; DB 2; Length 895;  
Best Local Similarity 99.5%; Pred. No. 2.4e-200;  
Matches 879; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 2358 AGAACCTCACTTCTTACTCTCTCAAAAGAGAGTGGGAAAGAACCATCAAACTTTCCT 2417  
Db 14 AGAACCTCACTTCTTACTCTCTCAAAAGAGAGTGGGAAAGAACCATCAAACTTTCCT 73

Qy 2418 CCTGACTTACCAACACAGGAAACAGACAGAGAGGGTGGCTCAGGACTTAGGACAGGGT 2477  
Db 74 CCTGACTTACCAACACAGGAAACAGACAGAGAGGGTGGCTCAGGACTTAGGACAGGGT 133

Qy 2478 ATAGCTTAGATGTGGAAAGCAAGAGGAGACGAGAGCTTGTAAATCACTGGCTAATGAGA 2537  
Db 134 ATAGCTTAGATGTGGAAAGCAAGAGGAGACGAGAGCTTGTAAATCACTGGCTAATGAGA 193

Qy 2538 AAAGGACACAGCTAACTCTAGGATGAAGCTGTGTACTAGGCTGGAGTTGCTTCTTGAAGA 2597  
Db 194 AAAGGACACAGCTAACTCTAGGATGAAGCTGTGTACTAGGCTGGAGTTGCTTCTTGAAGA 253

Qy 2598 TGGGACTCTTGGGTATCAAGACTTATGCCACATCACTGGGCTAGGGAAGTAGTGA 2657  
Db 254 TGGGACTCTTGGGTATCAAGACTTATGCCACATCACTGGGCTAGGGAAGTAGTGA 313

Qy 2658 TGGCAGCCCTCAAGCTGTCTTCAAGCAGGAGCTTGAAGAGTTATATTTGGCAGTGGCTC 2717  
Db 314 TGGCAGCCCTCAAGCTGTCTTCAAGCAGGAGCTTGAAGAGTTATATTTGGCAGTGGCTC 373

Qy 2718 CAATCTGTGACCAAGTATTTTCAAGCTTCTTCCCTGAAGATCAGGAGGGTGCCATTCATTGTC 2777  
Db 374 CAATCTGTGACCAAGTATTTTCAAGCTTCTTCCCTGAAGATCAGGAGGGTGCCATTCATTGTC 433

Qy 2778 TTTCTCTTAGCCCTCTAGGAAAGAGCTATTTTGTACTGTACCTTAGGGGTTCCT 2837  
Db 434 TTTCTCTTAGCCCTCTAGGAAAGAGCTATTTTGTACTGTACCTTAGGGGTTCCT 493

Qy 2838 GGAAGGAAACATGGAATCAGGATTTCTATAGCTGTAGTCCCTTCCACAGGGCCAT 2897  
Db 494 GGAAGGAAACATGGAATCAGGATTTCTATAGCTGTAGTCCCTTCCACAGGGCCAT 553

Qy 2898 GACTGGGAAAGGTATGGGACAGAGAGAAATTTGGGATTTTGGGTGCGAGCTACGCTCA 2957  
Db 554 GACTGGGAAAGGTATGGGACAGAGAGAAATTTGGGATTTTGGGTGCGAG-TACGCTCA 612

Qy 2958 CCCTAAACTTTTGTGGCTGGGCGATGTCTTGAAGCCGACACTGTGTAAGCAGGCTCTGC 3017  
Db 613 CCCTAAACTTTTGTGGCTGGGCGATGTCTTGAAGCCGACACTGTGTAAGCAGGCTCTGC 672

Qy 3018 TGGCTGTCTTACTCGTCCACACCTCTGACCTGTCTTGGAGACTCCATCCAGCCCCAG 3077  
Db 673 TGGCTGTCTTACTCGTCCACACCTCTGACCTGTCTTGGAGACTCCATCCAGCCCCAG 732

Qy 3078 GCAGCCACCTGTCTTCCAGCTCCACTATCTCCCTGTGAGCGGTGAACCTTCTGTACTG 3137  
Db 733 GCAGCCACCTGTCTTCCAGCTCCACTATCTCCCTGTGAGCGGTGAACCTTCTGTACTG 792

Qy 3138 TGTCTCGGTCATATATGAAATTTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTA 3197  
Db 793 TGTCTCGGTCATATATGAAATTTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTA 852

Qy 3198 CAAATAAAGATTATTTTCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3240  
Db 853 CAAATAAAGATTATTTTCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 895

RESULT 13  
AA27472  
ID AAS27472 standard; cDNA; 566 BP.  
AC AAS27472;  
XX  
DT 07-NOV-2001 (first entry)  
XX cDNA encoding novel signal transduction pathway protein, Seq ID 507.  
XX  
DE Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
XX anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;  
XX immune system disorder; rheumatoid arthritis; inflammatory condition;  
XX organ transplant rejection; infection; hepatitis C; blood disorder;  
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
XX reproductive system; gastrointestinal; liver disorder; AIDS; ss;  
XX acquired immune deficiency syndrome.  
XX Homo sapiens.  
XX  
XX WO200154733-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001312.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214986P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 14-JUL-2000; 2000US-0217496P.  
XX 26-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225366P.  
XX 14-AUG-2000; 2000US-0225367P.  
XX 14-AUG-2000; 2000US-0225368P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 14-AUG-2000; 2000US-0225759P.  
XX 18-AUG-2000; 2000US-0226279P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226686P.  
XX 22-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228242P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230437P.  
XX 06-SEP-2000; 2000US-0230438P.



Qy 182 GATGAGACCATAGAGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 241  
Db |||||||  
Qy 208 GATGAGACCATAGAGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 267  
Db |||||||  
Qy 242 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGCCCCAGGGGATATTTT 301  
Db |||||||  
Qy 268 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGCCCCAGGGGATATTTT 327  
Db |||||||  
Qy 302 GGTCTATGATATAGCAGCAGCGCTTTACACGACATCATGAAGTGGTCACTGACGT 361  
Db |||||||  
Qy 328 GGTCTATGATATAGCAGCAGCGCTTTACACGACATCATGAAGTGGTCACTGACGT 387  
Db |||||||  
Qy 362 GGATGAGTACGACCAAGCGCTTCAAGAGATCTTTATTTGGGAATAAGCTCATGAGGA 421  
Db |||||||  
Qy 388 GGATGAGTACGACCAAGCGCTTCAAGAGATCTTTATTTGGGAATAAGCTCATGAGGA 447  
Db |||||||  
Qy 422 GCAGAAACGGCAGGTGGGAAGAGAGAGAGCGCAGCAGCTCGCGAAGAGATATGGCATGGA 481  
Db |||||||  
Qy 448 GCAGAAACGGCAGGTGGGAAGAGAGAGAGCGCAGCAGCTCGCGAAGAGATATGGCATGGA 507  
Db |||||||  
Qy 482 CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCACTACGGGTCTGA 540  
Db CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCACTACGGGTCTGA 566

RESULT 14

ADB93650

ID ADB93650 standard; cDNA; 566 BP.

XX AC

XX ADB93650;

XX DT

XX 04-DEC-2003 (first entry)

XX XX

XX Human cDNA encoding a novel protein #497.

XX ss; Gene; human; autoimmune disease; Parkinson's disease; silicosis;  
KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;  
KW immunosuppressive agent; adjuvant; enhance immune response;  
KW higher affinity antibody induction;  
KW increased serum immunoglobulin concentration.

XX OS

XX Homo sapiens.

XX PN

XX US2002168711-A1.

XX XX

XX 14-NOV-2002.

XX PF

XX 17-JAN-2001; 2001US-00764868.

XX PR

XX 31-JAN-2000; 2000US-0179065P.

XX PR

XX 04-FEB-2000; 2000US-0180628P.

XX PR

XX 28-JUN-2000; 2000US-0214886P.

XX PR

XX 07-JUL-2000; 2000US-0216647P.

XX PR

XX 07-JUL-2000; 2000US-0216880P.

XX PR

XX 11-JUL-2000; 2000US-0217487P.

XX PR

XX 11-JUL-2000; 2000US-0217496P.

XX PR

XX 14-JUL-2000; 2000US-0218290P.

XX PR

XX 26-JUL-2000; 2000US-0220963P.

XX PR

XX 26-JUL-2000; 2000US-0220964P.

XX PR

XX 14-AUG-2000; 2000US-0224518P.

XX PR

XX 14-AUG-2000; 2000US-0224519P.

XX PR

XX 14-AUG-2000; 2000US-0225267P.

XX PR

XX 14-AUG-2000; 2000US-0225268P.

XX PR

XX 14-AUG-2000; 2000US-0225447P.

XX PR

XX 14-AUG-2000; 2000US-0225575P.

XX PR

XX 14-AUG-2000; 2000US-0225757P.

XX PR

XX 22-AUG-2000; 2000US-0226868P.

XX PR

XX 30-AUG-2000; 2000US-0228924P.

XX PR

XX 01-SEP-2000; 2000US-0229287P.

XX PR

XX 01-SEP-2000; 2000US-0229343P.

XX PR

XX 01-SEP-2000; 2000US-0229344P.

XX PR

XX 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244817P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

XX PI

XX WPI; 2003-719985/68.

XX DR

XX P-PSDB; ADB94263.

XX XX

XX New isolated polypeptide useful for diagnosing and treating  
PT immunosuppressive conditions such as autoimmune disease and Parkinson's  
PT disease.

XX Claim 3; SEQ ID NO 507; 345pp; English.

XX The invention relates to an isolated polypeptide. The polypeptide is

CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition in a subject, by determining the presence or  
CC amount of expression of the polypeptide in a biological sample and  
CC diagnosing a pathological condition or a susceptibility to a pathological  
CC condition based on the presence or amount of expression of the  
CC polypeptide. The polypeptide is also useful for identifying a binding  
CC partner to the polypeptide, which involves contacting the polypeptide  
CC with a binding partner and determining whether the binding partner  
CC effects an activity of the polypeptide. The polypeptide or the nucleic  
CC acid encoding the polypeptide is useful for preventing, treating, or  
CC ameliorating a medical condition, which involves administering the  
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
CC is useful for diagnosing a pathological condition or a susceptibility to  
CC a pathological condition in a subject, which involves determining the  
CC presence or absence of a mutation in the nucleic acid, and diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC based on the presence or absence of the mutation. The polypeptide, the  
CC nucleic acid and an antibody to the polypeptide are useful for treating  
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents cDNA encoding a novel human protein. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format direct from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20020168711.

XX Sequence 566 BP; 140 A; 145 C; 172 G; 103 T; 0 U; 6 Other;

XX SQ

Query Match 16.4%; Score 534.2; DB 9; Length 566;  
Best Local Similarity 99.9%; Pred. No. 1.2e-119;  
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCCCGTCGCCGCCGAGTCCCGCGCCCGCTGGCCCGCAGTCATGCGGAGCAGTACGA 61  
DB 28 GCCCGTCGCCGCCGAGTCCCGCGCCCGCTGGCCCGCAGTCATGCGGAGCAGTACGA 87  
QY 62 TGTCTCTGTCGCCGTCGTGTCATGCGGAGTCCCGCGCCCGCTGGCCCGCAGTACGA 121  
DB 88 TGTCTCTGTCGCCGTCGTGTCATGCGGAGTCCCGCGCCCGCTGGCCCGCAGTACGA 147  
QY 122 CCGCTTACCGAACAGATTCATCTCTGCGACATCTCCACCATCTCCACCATCTGACTTTAA 181  
DB 148 CCGCTTACCGAACAGATTCATCTCTGCGACATCTCCACCATCTCCACCATCTGACTTTAA 207  
QY 182 GATGAAGACCATAGAGGTAGAGCGGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 241  
DB 208 GATGAAGACCATAGAGGTAGAGCGGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 267  
QY 242 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGCGCGGCCCGAGGATATTTT 301  
DB 268 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGCGCGGCCCGAGGATATTTT 327  
QY 302 GGTCTATGACATTAGCAGCGAGCGCTCTTACAGACATCATGAAAGTGGTCACTGACGT 361  
DB 328 GGTCTATGACATTAGCAGCGAGCGCTCTTACAGACATCATGAAAGTGGTCACTGACGT 387  
QY 362 GGATGATACGACAGAGCGTCCAGAGATCTTATTTGGGATTAAGCTGATGAGGA 421  
DB 388 GGATGATACGACAGAGCGTCCAGAGATCTTATTTGGGATTAAGCTGATGAGGA 447  
QY 422 GCAGAAACGCGAGTGGAGAGAGAGAGGCGAGCAGCTGCGGAGGAGTATGGCATGGA 481  
DB 448 GCAGAAACGCGAGTGGAGAGAGAGAGGCGAGCAGCTGCGGAGGAGTATGGCATGGA 507  
QY 482 CTTCTATGAACAAGTGCCTGSCACCAACCTCAACATTAAAGAGTCATTACGCGTCTGA 540  
DB 508 CTTCTATGAACAAGTGCCTGSCACCAACCTCAACATTAAAGAGTCATTACGCGTCTGA 566

RESULT 15  
ABK44752  
ID ABK44752 standard; cDNA; 481 BP.

AC ABK44752;  
XX  
XX  
XX 05-JUN-2002 (first entry)

XX cDNA encoding colon tumour protein, SEQ ID No 303.  
XX Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;  
XX gene; ss.

XX Homo sapiens.  
XX OS  
XX WO200212328-A2.  
XX  
XX 14-FEB-2002.

XX 31-JUL-2001; 2001WO-US024218.  
XX  
XX 03-AUG-2000; 2000US-0223283P.  
XX 28-MAR-2001; 2001US-0279763P.  
XX 29-JUN-2001; 2001US-0302051P.

XX (CORI-) CORIXA CORP.  
XX  
XX King GE, Meagher MJ, Xu J, Secret H;  
XX WPI; 2002-241739/29.

XX New colon cancer polypeptides and polynucleotides, useful as vaccines,  
PT

PT for diagnosing, preventing, and treating colon cancer, and as markers for  
PT the progression of cancer.

XX Claim 1; SEQ ID NO 303; 147pp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.  
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical  
CC compositions, such as vaccines, for the diagnosis, prevention, and  
CC treatment of colon cancer. Polynucleotide sequences may be used as  
CC hybridisation probes or primers, and in the design and preparation of  
CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
CC proteins in tumour cells. The compositions are useful for stimulating an  
CC immune response against cancer, particularly for the immunotherapy of  
CC colon cancer, and as markers for the progression of cancer. ABK4450-  
CC ABK46237 represent coding sequences of human colon tumour proteins of the  
CC invention. Note: With the exception of SEQ ID No 1 and 2, the sequence  
CC data for this patent did not form part of the printed specification but  
CC was supplied by the European Patent Office

XX SQ Sequence 481 BP; 111 A; 113 C; 130 G; 125 T; 0 U; 2 Other;

Query Match 14.7%; Score 477.4; DB 6; Length 481;  
Best Local Similarity 99.4%; Pred. No. 7.7e-106;  
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 2641  
DB 1 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 60

QY 2642 CTAGGGAAGTAGGTATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTTGAGAAGTTA 2701  
DB 61 CTAGGGAAGTAGGTATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTTGAGAAGTTA 120

QY 2702 TATTGGGACGTGGCTCCAATCTGTGGACCAAGTATTTCCAGCTTTTCCCTGAAGATCAGG 2761  
DB 121 TATTGGGACGTGGCTCCAATCTGTGGACCAAGTATTTCCAGCTTTTCCCTGAAGATCAGG 180

QY 2762 GGTGCCATTCAATGTCTTTCTCTAGCCCTCAGGAAAGAGGACTATATTGTACT 2821  
DB 181 GGTGCCATTCAATGTCTTTCTCTAGCCCTCAGGAAAGAGGACTATATTGTACT 240

QY 2822 GTACCCCTAGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTATAGACTGATAGGCC 2881  
DB 241 GTACCCCTAGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTATAGACTGATAGGCC 300

QY 2882 TATCCACAAGGGCCATGACCTGGGAAAGGTATGGGACGAGAGAGAAATGGGATTTAG 2941  
DB 301 TATCCACAAGGGCCATGACCTGGGAAAGGTATGGGACGAGAGAGAAATGGGATTTAG 360

QY 2942 GGTGCAGCTACGCTACCCCTAAACTTTTGGTGGCTGGGCGCATGCTTTGAGGCCCCAGACT 3001  
DB 361 GGTGCAGCTACGCTACCCCTAAACTTTTGGTGGCTGGGCGCATGCTTTGAGGCCCCAGACT 420

QY 3002 GTTAAGCAGGCTGTGCTGGCTGTTTACTCGTACACCACTCTGCACTGCTGCTTGA 3061  
DB 421 GTTAACCCAGGCTGTGCTGGCTGTTTACTCGTACACCACTCTGCACTGCTGCTTGA 480

QY 3062 C 3062  
DB 481 C 481

Search completed: March 21, 2004, 05:37:11  
Job time : 1158 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2004, 23:06:10 ; Search time 226 Seconds  
(without alignments)  
7997.677 Million cell updates/sec

Title: US-09-817-198C-1  
Perfect score: 3257  
Sequence: 1 tgcgcgtgccgcgcgcag.....aaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

.Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.2	5.3	730	4	US-09-300-958A-16
2	156.6	4.8	1340	2	US-08-824-873-2
3	156.6	4.8	1340	3	US-09-198-184-2
4	156	4.8	925	2	US-08-916-901-4
5	156	4.8	925	4	US-09-154-602-4
6	144.6	4.4	3936	4	US-09-919-172-49
7	138	4.2	842	4	US-09-255-920A-6
8	136.2	4.2	875	4	US-09-075-454-10
9	136.2	4.2	1106	4	US-09-620-312D-959
10	136.2	4.2	2612	4	US-09-484-970B-142
11	133.4	4.1	1069	4	US-09-620-312D-646
12	124.2	3.8	723	4	US-09-016-434-1422
13	123.6	3.8	1102	4	US-09-620-312D-828
14	122.6	3.8	8137	4	US-09-566-921-7
15	120.6	3.7	847	2	US-08-773-423-4
16	115	3.5	833	4	US-09-620-312D-426
17	113	3.5	639	4	US-09-399-913-66
18	110	3.4	970	3	US-08-888-077A-28
19	106.2	3.3	1054	4	US-09-976-594-1096
20	102.2	3.1	3745	4	US-09-976-594-387
21	93.8	2.9	193	4	US-09-702-705-989
22	93.8	2.9	193	4	US-09-702-705-1161
23	93.8	2.9	193	4	US-09-736-457-989
24	93.8	2.9	193	4	US-09-736-457-1161
25	93.8	2.9	193	4	US-09-614-124B-989
26	93.8	2.9	193	4	US-09-614-124B-1161
27	93.8	2.9	193	4	US-09-671-325-989

28	93.8	2.9	193	4	US-09-671-325-1161	Sequence 1161, Ap
29	93.6	2.9	313	4	US-09-313-294A-7460	Sequence 7460, Ap
30	91.2	2.8	820	3	US-08-741-411-6	Sequence 6, Appli
31	91.2	2.8	820	4	US-09-016-434-112	Sequence 112, App
32	85.8	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl
33	84	2.6	241	4	US-09-016-434-813	Sequence 813, App
34	81.2	2.5	921	4	US-09-016-434-1124	Sequence 1124, Ap
35	80.6	2.5	890	3	US-08-741-411-4	Sequence 4, Appli
36	80	2.5	1172	4	US-09-075-454-8	Sequence 8, Appli
37	80	2.5	1255	2	US-08-766-551-6	Sequence 6, Appli
38	80	2.5	1533	4	US-09-075-454-11	Sequence 11, Appl
39	80	2.5	1886	4	US-09-620-312D-647	Sequence 647, App
40	79.2	2.4	1984	4	US-09-023-655-7	Sequence 7, Appli
41	79	2.4	848	3	US-08-741-411-2	Sequence 2, Appli
42	78.8	2.4	912	4	US-09-016-434-1159	Sequence 1159, Ap
43	77.4	2.4	503	4	US-09-621-976-15288	Sequence 15288, A
44	77	2.4	1749	4	US-09-149-476-54	Sequence 54, Appl
45	75.6	2.3	1407	4	US-09-493-914-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-300-958A-16  
; Sequence 16, Application US/09300958A  
; Patent No. 6495319  
; GENERAL INFORMATION:  
; APPLICANT: McClelland, Michael  
; APPLICANT: Welsh, John  
; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of  
; TITLE OF INVENTION: Using Same  
; FILE REFERENCE: P-PH 3457  
; CURRENT APPLICATION NUMBER: US/09/300,958A  
; CURRENT FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/083,331  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/098,070  
; PRIOR FILING DATE: 1998-08-27  
; PRIOR APPLICATION NUMBER: 60/118,624  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 730  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-300-958A-16

Query Match		5.3%;	Score 174.2;	DB 4;	Length 730;
Best Local Similarity		55.7%;	Pred. No. 4.9e-36;	Indels 3;	Gaps 1;
Matches 342;		Conservative 0;	Mismatches 258;		
Qy	7	CTGCCCCCGCCGAGTTCCTCCGCGCCCGCGTGGCCCCAGTCATCGCGAAGCAGTACGATGTGC	66		
Db	44	CGGACCCCGCCCGCGCGCGCTCCGCCCGCCCGCCATCGCGCGGACTACGACCACC	103		
Qy	67	TGTTCCCGGCTGCTGCTGATCGGGGACTCCCGGGTGGGCAAGACCTGCTGTCGCCCT	126		
Db	104	TCTTCAAGCTGCTCATCATCGGCACAGCGGTGTGGCAAGCAGTCTTACTGTTCGTT	163		
Qy	127	TCACCGACAGGAGTTCACCTCTCCGACATCTCCACCATCGGTGTTGACTTTAAGATGA	186		
Db	164	TTGAGACAAACACTTTCTTCAGGCGAGTACATCACACATCGGAGTGGATTTCAAGATCC	223		
Qy	187	AGACATAGAGGTAGAGCGGCATCAAGTGGCGATACAGATCTGGGACACTCGAGGCAGG	246		
Db	224	GGACCGTGGAGATCAACGGGGAGGTGAAGCTGCAGATCTGGGACACACACGGGGCAGG	283		
Qy	247	AGAGATACAGACCATCAAAAGCAGTACTATCGCGGGGCCACAGGGGATATTTTGTGCT	306		
Db	284	AGCGTTCCGACCATCACTCCACGTATTATCGGGGGACCCACGGGGTCTATTGTGTTT	343		



307 ATGACATTACGAGCGGCTCTTACAGACATCATGAAGTGGTTCAGTGCCTGATG 366  
344 ACAGCGTCACCGAGTCCGAGTCTCTTGTCAACGTCAGCGGTGGCTTCACGAATCAACC 403  
367 AGTACGACACGAGCGGCTCCAGAGATCTTATTGGGAATAAGGCTGATGAGGAGCAGA 426  
404 AGAAC-----TGTGATGATGTGTGCGGAATATTAGTGGGTAAATAGATGACGACCCCTGAGC 460  
427 AACCGCAGGTGGGAAGAGAGACGAGCGGCGAGCTGGCGGAAGAGATATGGCATGGACTTCT 486  
461 GGAAGTGTGGAGACGGAAGATGCCATCAAAATTCGCCGGCGAGATGGGCATCCAGTTGT 520  
487 ATGAAACAAGTGCCTGACCAACCTCAACATTAAGAGTCAATTCACCGCTCTGACAGAGC 546  
521 TCGAGACGAGCGCCACAGGAGAAATGTCAACGTGGAAGAGATGTTCACTGTCATCACGAGC 580  
547 TGGTCTGCGAGGCCCATAGGAAGAGCTGGAAGGCTCCGAGATGCGTGGCGAGCAATGAT 606  
581 TGGTCTCCGAGCAAGAAAGACAACTGGCAAAACAGCAGCAGCAACACAGAACGATG 640  
607 TGG 609  
641 TGG 643

RESULT 2  
US-08-824-873-2  
; Sequence 2, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,873  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1340 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT04  
; CLONE: 738957  
US-08-824-873-2

Query Match 4.8%; Score 156.6; DB 2; Length 1340;

Best Local Similarity 59.0%; Pred. No. 3e-31;  
Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2;

QY 57 TACGATGTCTTTCGGGCTGTCTGATCGGGGACTCGGGGTGGGCAAGACCTGGCTG 116  
Db 22 TACGACGTCCGCTTCAAGGTCTATGCTGTGGGGGACTCGGGGTGGGCAAGACCTGTCTG 81  
QY 117 CTGTGC-CGCTTTCACGCAACGAGTTCACATCTCTCGCACATCTCCACCATCGGTG--TT 173  
Db 82 CTGGGTGGCATTCAGGATGGTGTCTTCTGCGGGGACCTTCATCTCCACCGTAGCAT 141  
QY 174 GACTTTTAAGATGAAGACCATAGAGGTAGAGCGGCATCAAAAGTCCGGATACAGATCTGGGAC 233  
Db 142 GACTTCCGGAACAAAGTTCTGACGCTGGATGTTGTGAAGGTGAAGCTGCAGATGTGGGAC 201  
QY 234 ACTGAGGCGAGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGGCCCCAGGG 293  
Db 202 ACAGCTGTGTGAGGCGGTTCGCGCATGTGTACCCATGCTTACCGGATGCTCATGCT 261  
QY 294 ATATTTTTCGTCTATGACATTAGCAGCGCTCTTACGACCATCATGAAGTGGGTC 353  
Db 262 CTGCTGCTCTTACGATGTCAACAAAGGCTCTCTTTGACAAACATCCAGGCTGGCTG 321  
QY 354 AGTGACGTGGATGATGACGACCAAGCGCTGCGCAAGAGATCTTATTTGGGAATAAGGCT 413  
Db 322 ACCGAGATCCAGGATACGCCAGCACGACGCTGGCGCTCATGCTGTGGGGAACAAAGGTG 381  
QY 414 GATGAGGAGCAGNAACGCGCAGGTGGGAAGAGCAGAGGCGCAGCTGGCGAAGGAGTAT 473  
Db 382 GACTCTGCCCATGAGCGTGTGTGAAGAGGGAGGACGGGGAGAGCTGGCCAAAGGAGTAT 441  
QY 474 GGCATGGACTTCTATGAACAAAGTGCCTGCACCACTCAACATTAAGAGTCAATTCAGC 533  
Db 442 GGAATGCCCTTCTATGAGACCGCGCCAAAGCGGCTCAACGTGGACTTGGCTTTCACA 501  
QY 534 CGTCTGACAGAGCTGTGTGCTGAGGCGCCCATPAGGAAG 570  
Db 502 GCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGG 538

RESULT 3  
US-09-198-184-2  
; Sequence 2, Application US/09198184  
; Patent No. 6010859  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/198,184  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,873  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1340 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT04  
; CLONE: 738957  
US-09-198-184-2

TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1340 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PANCNOT04  
CLONE: 738957  
US-09-198-184-2

Query Match 4.8%; Score 156.6; DB 3; Length 1340;  
Best Local Similarity 59.0%; Pred. No. 3e-31;  
Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2;

QY 57 TACGATGCTGTTCCGGCTGCTGCTGATCGGGGACTCGGGGTGGGCAAGACCTGCTG 116  
Db |||||  
QY 22 TACGACGTCGCTTCAAGGTCAATGCTGTGGGGGACTCGGGTGGGGAAGACCTGCTG 81  
Db |||||  
QY 117 CTGTGC-CGCTTCCCGCAACAGGTTCCTCTCGGCACATCTCCACCATCGGTG--TT 173  
Db |||||  
QY 82 CTGGTGGATCAAGGATGCTGCTTCTCGGGGGACCTTCATCTCCACCGTAGCAT 141  
Db |||||  
QY 174 GACTTTAAGATCAAGACCATAGAGGTAGACGGGCATCAAGGTGCGGATACAGATCTGGGAC 233  
Db |||||  
QY 142 GACTTCCGGACAAAGTTCTGACGTGATGTTGTAAGGTGAAGCTGCAGATGTTGGAC 201  
QY 234 ACTGAGGGGAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCGAGGG 293  
Db |||||  
QY 202 ACAGCTGTGAGGAGCGGTTCGCGAGTGTACCCATGCTACTACCGGGATGCTCATGCT 261  
QY 294 ATATTTTGGTCTATGACATAGAGCGAGCGCTTACACGACATCATGAAGTGGGTC 353  
Db |||||  
QY 262 CTGCTGCTCTAGATGTACCAACAAAGGCTCTTTTGCAACATCCAGGCTTGGCTG 321  
QY 354 AGTGACGTGGATGATGACGACCAAGGCGTCCAGAAAGATCTTATTGGGAATAAGGCT 413  
Db |||||  
QY 322 ACCGAGATCCAGAGTACGCCGACGACGACGCTGGCGCTCATGCTGCTGGGAAACAGGTG 381  
QY 414 GATGAGGAGCAAAACGCGAGGTGGGAAAGAGAGCAAGGCGAGCTGGCGAAGAGTAT 473  
Db |||||  
QY 382 GACTCTGCCATGAGCGTGTGTGAAGAGGAGGACGCGGAGAGCTGGCCAAAGAGTAT 441  
QY 474 GGCATGGACTTCTATGAACAGTGCCTGACCAACCTCAACATTAAGAGTCAATCAAG 533  
Db |||||  
QY 442 GGACTGCCCTTCATGGAGACGAGCGCCCAAGACGGGCTCAACGCTGAGCTTGGCCTTACA 501  
QY 534 CGTCTGACAGAGCTGGTCTGCGAGCCCATAGGAAGG 570  
Db |||||  
QY 502 GCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGG 538

RESULT 4  
US-09-916-901-4  
; Sequence 4, Application US/08916901  
; Patent No. 5892012  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA: US/08/916,901  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 925 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVRUT04  
CLONE: 2514506  
US-08-916-901-4

Query Match 4.8%; Score 156; DB 2; Length 925;  
Best Local Similarity 56.7%; Pred. No. 3.5e-31;  
Matches 288; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 36 GCCCAGCTGTCGGGAAGCAGTACGATGCTGTTCCGGCTGCTGCTGATCGGGGACTCC 95  
Db |||||  
QY 57 GCCGCCCATGAACCCCGAATATGACTACTGTTAAGTGTCTTTTGAATGGGACTCA 116  
Db |||||  
QY 96 GGGGTGGGCAAGACTGCTGCTGTCGCGTTCACCGAACAGAGTTCACTCTCTCGCAC 155  
Db |||||  
QY 117 GCGGTGGGCAAGTCACTGCTGCTCTGCTGATGACAGTACACAGAGAGCTAC 176  
QY 156 ATCTCCACCATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTG 215  
Db |||||  
QY 177 ATCAGCACCATCGGGGTGGACTTCAAGATCCGAACCATCGAGCTGGATGGCAAACTATC 236  
QY 216 CGGATACAGATCTCGGACACTGCGAGGCGAGAGATACAGACCATCACAAAGCAGTAC 275  
Db |||||  
QY 237 AACCTCAGATCTGGGACACAGCGGGCCAGAACGGTTCGGGACCATCACTTCAGCTAC 296  
QY 276 TATCGGGGGCCCGGGGATATTTTGGTCTATGACATTAGCAGGAGCGCTCTTTACAG 335  
Db |||||  
QY 297 TACCGGGGGCTCATGGCATCATCGTGTGTATGACGCTCACTGACAGGAATCTTACGCC 356  
QY 336 CACATCATGAAGTGGGTCACTGAGCTGGATGAGTACGACCAAGAGGCTCCAGAGATC 395  
Db |||||  
QY 357 AACGTGAAGCAGTGGCTCGAGGATTTACCGCTATGCGAGGAGAACGTCAATAAGCTC 416  
QY 396 CTATTTCGGATAGGCTGATGAGGACGAGAACCGGAGTGGGAGAGAGCAGAGGGCAG 455  
Db |||||  
QY 417 CTGGTGGGCAACAAGAGCGACCTCACCCAAAGAGTGGTGGCAACACACAGCCCAAG 476  
QY 456 CAGCTGGCAGGAGTATGCGATGGAATCTTATGAACAAGTGCCTGCACCAACTCAAC 515  
Db |||||  
QY 477 GAGTTTCAGACTCTCTGGGATCCCTCTTTGAGAGGAGCGGCCAAGATGCCACCAAT 536  
QY 516 ATTAAGAGTCAATTCACGCGTCTGACAG 543  
Db |||||  
QY 537 GTCGAGCAGCGTTTCATGACCATGGCTG 564

RESULT 5  
US-09-154-602-4  
; Sequence 4, Application US/09154602  
; Patent No. 6300472



```

US-09-255-920A-6
; Sequence 6, Application US/09255920A
; Patent No. 6623980
; GENERAL INFORMATION:
; APPLICANT: Fisher, Joseph
; APPLICANT: Lorens, James
; APPLICANT: Anderson, David
; APPLICANT: Luo, Ying
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS
; FILE REFERENCE: A65905-1/DJB/RMS
; CURRENT APPLICATION NUMBER: US/09/255,920A
; CURRENT FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 60/075,534
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: 60/086,650
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3)
; OTHER INFORMATION: The n at position 3 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (6)..(8)
; OTHER INFORMATION: The n at positions 6 through 8 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (12)
; OTHER INFORMATION: The n at position 12 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (14)..(15)
; OTHER INFORMATION: The n at positions 14 through 15 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (17)..(18)
; OTHER INFORMATION: The n at positions 17 through 18 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (22)
; OTHER INFORMATION: The n at position 22 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (25)
; OTHER INFORMATION: The n at position 25 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (726)
; OTHER INFORMATION: The n at position 726 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (823)
; OTHER INFORMATION: The n at position 823 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (108)
; OTHER INFORMATION: The n at position 108 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (138)
; OTHER INFORMATION: The n at position 138 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (204)
; OTHER INFORMATION: The n at position 204 represents an unknown.
; OTHER INFORMATION: The n at position 204 represents an unknown.
US-09-255-920A-6

Query Match          4.2%; Score 138; DB 4; Length 842;
Best Local Similarity 57.6%; Pred. No. 1.9e-26;
Matches 265; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

RESULT 7
US-09-255-920A-6
; Sequence 6, Application US/09255920A
; Patent No. 6623980
; GENERAL INFORMATION:
; APPLICANT: Fisher, Joseph
; APPLICANT: Lorens, James
; APPLICANT: Anderson, David
; APPLICANT: Luo, Ying
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS
; FILE REFERENCE: A65905-1/DJB/RMS
; CURRENT APPLICATION NUMBER: US/09/255,920A
; CURRENT FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 60/075,534
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: 60/086,650
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3)
; OTHER INFORMATION: The n at position 3 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (6)..(8)
; OTHER INFORMATION: The n at positions 6 through 8 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (12)
; OTHER INFORMATION: The n at position 12 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (14)..(15)
; OTHER INFORMATION: The n at positions 14 through 15 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (17)..(18)
; OTHER INFORMATION: The n at positions 17 through 18 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (22)
; OTHER INFORMATION: The n at position 22 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (25)
; OTHER INFORMATION: The n at position 25 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (726)
; OTHER INFORMATION: The n at position 726 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (823)
; OTHER INFORMATION: The n at position 823 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (108)
; OTHER INFORMATION: The n at position 108 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (138)
; OTHER INFORMATION: The n at position 138 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (204)
; OTHER INFORMATION: The n at position 204 represents an unknown.
; OTHER INFORMATION: The n at position 204 represents an unknown.
US-09-255-920A-6

; 73 GGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTCTGCTGTGCGCTTACCG 132
; 119 GGGTGATGCTTCTTGGAGANTCGGGCGTGGCAAAACCTTTCTCTGATCCAAATTAAG 178
; 133 ACAACGAGTTCCACTCCT---CGCACATCTCCACCATCGGTGTTGACTTTTAAGATGAGA 189
; 179 ACGGGGCTTCTCTGTCGGGAACCTTATAGCCACCGTCGGCATAGACTTCAGGAATAAAG 238
; 190 CCATAGAGGTAGACGGCATCAAGTGGCGGATACAGATCTGGGACACTGCGAGGAGGAGA 249
; 239 TGGTGACAGTGGATGGTTCAGGGTGAAGCTTCAGATCTGGGACACTGCGAGCAGGAGC 298
; 250 GATACAGAGCATCAAAAGCAGTACTATCGGGGGCCCGAGGAGATTTTTTGGTCTATG 309
; 299 GCTTCCGAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCTGTTGTATG 358
; 310 ACATTAGCAGGAGCGCTCTTACCAGCAGATCAAGTGGGTGGTGGTGGTGGTGGTGGT 369
; 359 ACATCAACCAACAGTCTCTTTTGACAAACATCAGGGCTGGCTCACAGAGATTATGAGT 418
; 370 AGCACCAGAGAGGCTCCAGAAGATCCTTATTGGGAATAAGGCTGTAGGAGGAGAAAC 429
; 419 ATGCCAGAGGAGCTGGTGTATGTTCTTAGGCAACAGGCCGATGTAGAGCGGAA 478
; 430 GGCAGTGGGAAGAGAGCAAGGGCAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATG 489
; 479 GGGTGATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAATATGGTGTCTCTTCATGG 538
; 490 AAACAAGTGGCTGCACCAACCTCAACATTAAAGATCATTT 529
; 539 AGACAGTGGCAAGACTGGGCATGAACGTGGAGTTGGCCTT 578

US-09-075-454-10
; Sequence 10, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
```

TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 875 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: UCMCL5T01  
CLONE: 1528559  
US-09-075-454-10

Query Match 4.2%; Score 136.2; DB 4; Length 875;  
Best Local Similarity 57.5%; Pred. No. 5.7e-26;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

Qy 57 TACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116  
Db 72 TACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTC 131  
Qy 117 CTGTGCGCTTCCCGCAACAGGTTCCACTCCT---CGCACATCTCCACCATCGGTGTT 173  
Db 132 CTGATCCAAATTCAAAGACGGGGCTTCTCTCGGAACCTTCATAGCCACCGTCGGCATA 191  
Qy 174 GACTTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTGCAGATACAGATCTGGGAC 233  
Db 192 GACTTCAGAACAGGTGGTGACTGTGTGATGGCTGAGAGTGAAGCTGCAGATCTGGGAC 251  
Qy 234 ACTGAGGCGAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGGCCAGGGG 293  
Db 252 ACCGCTGGCAGGAACGGTTCGGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGGCC 311  
Qy 294 ATATTTTGGTCTATGACATTAGCAGCGGCTCTTACCAGACATCATGAGTGGGTC 353  
Db 312 TTGCTTCTGTGTATGATCATCCAAACAATCTTCTTCGACAACATCAGGCGCTGGCTC 371  
Qy 354 AGTGACGTGATGATGACGACCAAGAGCGCTCCAGAGATCTTATTGGGAATAAGGCT 413  
Db 372 ACTGAGATTCATGATGATGATGCCAGAGGACGTGTGTATGCTGTAGGCAACAAGCG 431  
Qy 414 GATGAGAGCAAGAACGGCAGGTGGGAGAGAGCAAGGCGAGCAGTGGGGAAGAGTAT 473  
Db 432 GATATGAGCAGCAAGAGTATCGCTTCCGAAGACGAGAGACCTTGGCCAGGGAGTAC 491  
Qy 474 GGCATGACTTCTATGAACAAGTGCCTGACCAACCTCAA 514  
Db 492 GGTGTTCCCTTCTGGAGACGAGCGCAAGACTGGCATGAA 532

## RESULT 9

US-09-620-312D-959  
; Sequence 959, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenchua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yuning  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 959  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (124)..(699)  
; US-09-620-312D-959

Query Match 4.2%; Score 136.2; DB 4; Length 1106;  
Best Local Similarity 57.5%; Pred. No. 6.5e-26;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

Qy 57 TACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116  
Db 103 TACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTC 162  
Qy 117 CTGTGCGCTTCCCGCAACAGGTTCCACTCCT---CGCACATCTCCACCATCGGTGTT 173  
Db 163 CTGATCCAAATTCAAAGACGGGGCTTCTCTCGGAACCTTCATAGCCACCGTCGGCATA 222  
Qy 174 GACTTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAGTGCAGATACAGATCTGGGAC 233  
Db 223 GACTTCAGGAACAAGGTGGTGACTGTGTGATGGCGTGAGAGTGAAGCTGCAGATCTGGGAC 282  
Qy 234 ACTGAGGCGAGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGGCCAGGGG 293  
Db 283 ACCGCTGGGCGAGGAACGGTTCGGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCC 342  
Qy 294 ATATTTTGGTCTATGACATTAGCAGCGGCTCTTACCAGACATCATGAGTGGGTC 353  
Db 343 TTGCTTCTGTGTATGACATCAACAATCTTCTTCGACAACATCAGGCGCTGGCTC 402  
Qy 354 AGTGACGTGATGATGACGACCAAGAGCGCTCCAGAGATCTCTTATTGGGAATAAGGCT 413  
Db 403 ACTGAGATTCATGATATGCCAGAGGAGCTGTGTATGCTGTCTAGGCAACAAGCG 462  
Qy 414 GATGAGAGCAGAAACCGGAGGTGGGAGAGAGAGGCGAGCGCTGGCGAAGAGTAT 473  
Db 463 GATATGAGCAGCAAGAGTATCGCTTCCGAAGACGAGAGACCTTGGCCAGGGAGTAC 522  
Qy 474 GGCATGACTTCTATGAACAAGTGCCTGACCAACCTCAA 514  
Db 523 GGTGTTCCCTTCTGGAGACGAGCGCAAGACTGGCATGAA 563

## RESULT 10

US-09-484-970B-142  
; Sequence 142, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkuth, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 142  
; LENGTH: 2612  
; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1  
US-09-484-970B-142

Query Match  
Best Local Similarity 4.2%; Score 136.2; DB 4; Length 2612;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 57 TACGATGCTGTTCCGGCTCTCTGATCGGGAGCTCCGGGTGGGCAAGACTGCTGCT 116  
DB 91 TACGACCTCAGCGGCAAGGTGATGCTCTGGGAGACACAGCGCTGGGCAAAACATGTTT 150  
QY 117 CTGTGCGCTTACCGCAACAGTTCACCTCTCT---CGCACATCTCCACCATCGGTGTT 173  
DB 151 CTGATCCATCAAGACGGGCTCTCTGTCGGAACTTCATAGCCACCGTCGGCATA 210  
QY 174 GACTTTAAGATGAAGCATAGAGGTAGACGGCATCAAGTGGCGATACAGATCTGGGAC 233  
DB 211 GACTTCAGGAACAAGGTGCTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGAC 270  
QY 234 ACTGCGGCGAGGAGATACACAGACCATCAAAAGCAGTACTATCGCGGGCCAGGGG 293  
DB 271 ACCCTGGCGAGGACGGTTCGGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGGCC 330  
QY 294 ATATTTTGGTCTATGACATTAGCAGCGAGCGCTTTACAGACATCATGAAGTGGTC 353  
DB 331 TTGCTTCTGTATGATCATCAACACATCTCTTTTCGACACATCAGGCGCTGCTC 390  
QY 354 AGTGACGTGGATGATGACGACCAAGAGCGGCTCCAGAGATCTCTTTATGGGAATPAAGGCT 413  
DB 391 ACTGAGATTCATGATGATGCTCCAGAGGAGCGTGTGATCATGCTGTAGGCAACAAGGCG 450  
QY 414 GATGAGAGCAAGAACGCGAGGTGGAGAGAGAGCAAGCGAGCTGGCGAAGAGTAT 473  
DB 451 GATATGAGCAGCAAGAGTATCGCTTCCGAAGACGAGAGACCTTTGGCGAGGAGTAC 510  
QY 474 GGCATGACTTCTATGAACAAAGTGCCTGACCAACCTCAA 514  
DB 511 GGTGTTCCCTTCTGGAGACCGCGCCAAAGACTGGCATGAA 551

## RESULT 11

US-09-620-312D-646  
Sequence 646, Application US/09620312D  
Patent No. 659662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunqing  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 659662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 646  
LENGTH: 1069  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (156)...(839)  
US-09-620-312D-646

Query Match  
Best Local Similarity 4.1%; Score 133.4; DB 4; Length 1069;  
Matches 266; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 52 AGCAGTACGATGCTGCTCCGGCTGCTGATCGGGAGCTCCGGGTGGGCAAGACCT 111  
DB 229 AGAATTTTGTACTCATGTTCAAATTTACTCATCGGCAATAGCAGTGTGGGAAACAT 288  
QY 112 GCCTGCTGTCCGCTTCAACGCAACAGGTTCCACTCTCCGACATCTCCACCATCGGTG 171  
DB 289 CTTTTCTATTCGGTATGAGATGACTCTCTTTACATCTGCTATCGTCAGCAGTTGGGA 348  
QY 172 TTGACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTGGCGATACAGATCTGGG 231  
DB 349 TCGATTTCAAAGTAAAACTGTATTCAAAAATGAAAAGAGAAATCAAGCTTCAGATTTGGG 408  
QY 232 ACATGCGGCGAGGAGATACAGACCATCAACAGCAGTACTATCGGGGGCCCCAGG 291  
DB 409 ACACGAGCGCGGAGGAAAGATACAGGACTATCACACAGCTTATTCGTGGAGCCATGG 468  
QY 292 GGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGGG 351  
DB 469 GCTTTATTTAATGATGATGACATTAACAATGAGAAATCCTTCAATGCGAGTACAGATTGGT 528  
QY 352 TCAGTACGCTGGATGATGACGACCAAGAGCGGCTCCAGAGATCTTTATTTGGGAATTAAG 411  
DB 529 CAATCAATCAAAACATACCTCTTGGGACAAATGCCCAAGTTATTCTGTTGGGAACAAGT 588  
QY 412 CTGATGAGGAGCAGAAACGCGAGGTGGAGAGAGAGCAAGGCGAGCAGCTGGCGAAGGAT 471  
DB 589 GTGACATGGAAGACGAGCGGCTCATCTCAACTGAGCGAGGTCAACATTTTAGGAGAACAGC 648  
QY 472 ATGGCATGGACTTCTATGAACAAAGTGCCTGACCAACCTCAACATTTAAAGAGTCAATCA 531  
DB 649 TTGGTTTGGATTTTGTGAAACAAAGTCCCAAGGACAAATTAATGTCAGCAGACATTTG 708  
QY 532 CGCGTCT 538  
DB 709 AGCGCT 715

## RESULT 12

US-09-016-434-1422  
Sequence 1422, Application US/09016434  
Patent No. 650938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible





Qy 492 ACAAGTGCCTGCACCAACCTCAACATTAA---AGAGTCATTACAGCGTCTGACAGAGCTG 548  
Db |||||  
693 ACCTCAGCCCTGGACTCTACCAATGTTGAGCTAGCTCTTGGAGAAATC 752  
Qy 549 GTGCTCAGAGCCCATAGAGAGAGCTGGAAGCGCTCCGGATGCGTGCAGCAAT 602  
Db |||||  
753 TTTCCGAAGGTGTCCAGCAGAGACAGAACAGCATCGCGACCAATGCCATCACT 806

RESULT 14  
US-09-566-921-7  
; Sequence 7, Application US/09566921  
; Patent No. 6682888  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE  
; FILE REFERENCE: PA-0024 US  
; CURRENT APPLICATION NUMBER: US/09/566,921  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 8137  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6682888 411474.10  
; NAME/KEY: unsure  
; LOCATION: 3488-3788  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-566-921-7

Query Match 3.8%; Score 122.6; DB 4; Length 8137;  
Best Local Similarity 53.2%; Pred. No. 7.7e-22;  
Matches 260; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
Qy 55 AGTAGATGCTGCTTCGGCTGCTGCTGATCGGGAGCTCGGGGTGGGGAAGACCTGCC 114  
Db |||||  
222 AATATGATTATTATTCAAGTTACTTCTGATGGCGACTCAGGGGTTGGAAAGTCTTGCC 281  
Qy 115 TGCTGTGCGGCTTCACGCACACAGAGTTCCTACTCTCGCACATCTCCACCATCGGTGG 174  
Db |||||  
282 TCTCTTAGTGTGCGATGATACATATACAGAAAGCTACATGACACAATGGGTGG 341  
Qy 175 ACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACA 234  
Db |||||  
342 ATTTCAAATAAAGAACTATAGATTAGACGGGAAACAACTCAAGCTTCAATATGGGACA 401  
Qy 235 CTGAGGCGCAGGAGATACAGACCATCAAAAGAGTACTATCGGGGGGCCAGGGGA 294  
Db |||||  
402 CAGCAGGCGCAGGAAAGATTTCGAACAATCACCTCCAGTTATTACAGAGAGCCCATGSCA 461  
Qy 295 TATTTTGGCTGTAGACATTAGCAGCGCGCTCTTACCAGACATCATGAAGTGGGTCA 354  
Db |||||  
462 TCATAGTTGTGTATGATGTACAGATCAGGAGTCTCTCAATAATGTTAAACAGTGGCTGC 521  
Qy 355 GTGACGTGGATGAGTACGACCAAGCGCTCCAGAAGATCCTTATTGGGAATAAGGCTG 414  
Db |||||  
522 AGGAAATAGATCGTTATGCCAGTGAATAATGTCACAAATTTGTTAGGGAACAAATGTG 581  
Qy 415 ATGAGGACAGAAACGGCAGGTGGGAGAGAGACGAGGCGCAGCTGGCGAAGGATG 474  
Db |||||  
582 ATCTGACCACAAAGAAAGTAGTAGACTACACAACAGCGGAAGGAATTTGCTGATTCCTTG 641  
Qy 475 GCATGGACTTCTATGAACAGTGCCTGCCAACCACTCAACATTAAAGAGTCATTACGC 534  
Db |||||  
642 GAAATCCGCTTTTGGAAACCAAGTGTGAAGATGCAACGAATGTAGAACAGTCTTTTCATGA 701  
Qy 535 GTCTGACAG 543

Db 702 CGATGGCAG 710  
|||

RESULT 15  
US-08-773-423-4  
; Sequence 4, Application US/08773423  
; Patent No. 5869291  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Bandman, Olga  
; TITLE OF INVENTION: NOVEL RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/773,423  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0183 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 847 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Consensus  
; CLONE: Consensus  
US-08-773-423-4

Query Match 3.7%; Score 120.6; DB 2; Length 847;  
Best Local Similarity 51.5%; Pred. No. 7.3e-22;  
Matches 301; Conservative 0; Mismatches 280; Indels 3; Gaps 1;  
Qy 22 TCCCGGCGCGCTCGCCCGCATGCGGAGCAGTAGTGCTGTTCGGGTGCTGC 81  
Db |||||  
60 TCGGAGCAAGATGGGGAATGGAATCTGAGGAGATTAATCTTGTCTTCAAGGTGCTGC 119  
Qy 82 TGATCGGGGAGCTCGGGGTGGGCAAGACCTGCTGTGTCGGCTTCCACGACAAACGAGT 141  
Db |||||  
120 TGATCGCGCAATCAGGTGTGGGGAAGACCAATCTACTCTCCGATTCACGCGCAATGAGT 179  
Qy 142 TCCACTCTCGCATCTCCACCATCGGTGTGTGATCTTTAAGATGAAGACCATAGAGGTAG 201  
Db |||||  
180 TCAGCCACGACAGCGCACCAATCGGGGTGTGAGTTCTCCACCGGCACCTGTGTGTTGG 239  
Qy 202 ACGGCATCAAGTCCGGATACAGATCTGGGACATCTGCGAGGCGCAGGAGATACCAAGCCA 261  
Db |||||  
240 GCACCGTGTGTCAAGGCTCAGATCTGGGACACAGCTGGCTGGAGGGTACCGAGCCA 299  
Qy 262 TCACAAAGCAGTACTATCTCGCGGGGCCCGGGGATATTTTGGTCTATGACATTAGCAGCG 321

Db	300	TCACCTCGGCGTACTATCGTGGTGCAGTGGGGCCCTCCTGGTGTGTGACCTAACCAAGC	359
Qy	322	AGCGCTTTACCAAGCACAATCATGAAGTGGGTGAGTGGTGGATGAGTACGACCAAG	381
Db	360	ACCAGACCTATGCTGTGGTGGAGCGATGGCTGAAGGAGCTCTATGACCATGCTGAAGCCA	419
Qy	382	CGGTCCAGAAGATCCTTATTGGGNAATGAGCTGATGAGGAGCAGAACGGCAGGTGGGA	441
Db	420	CGATCGTTCGTATGCTCGTGGGTAAACAAAAGTGACCTCAGCCAGGGCCGGGAAGTGCCCA	479
Qy	442	GAGAGCAAGGCGACAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAAACAAGTGCCT	501
Db	480	CTGAGGAGGCCGGAATGTTGCTGAAAACAATGGACTGCTTCTCTGGAGACCTCAGCCC	539
Qy	502	GCACCAACCTCAACATTAA---AGAGTCATTCACGCGTCTGACAGAGCTGGTGTGAGG	558
Db	540	TGGACTCTACCAATGTTGAGCTAGCCCTTGAGACTGTCTCTGAAAGAAATCTTTGCCAAGG	599
Qy	559	CCCATAGGAGGAGCTGGAAGGCCCTCCGGATGCGTGCCAGCAAT	602
Db	600	TGTCCAAGCAGAGACAGAACAGCATCCGGACCAATGCCATCACT	643

Search completed: March 21, 2004, 02:20:55  
Job time : 231 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 02:05:45 ; Search time 1083 Seconds  
(without alignments)  
11124.786 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcccgctgccgcccgcag.....aaaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3257	100.0	3257	9	US-09-817-198A-1
2	2696.6	82.8	28770	9	US-09-817-198A-3
3	1192.8	36.6	7924	14	US-10-311-455-2111
4	1192.8	36.6	7924	16	US-10-257-166-151
5	1160.8	35.6	7924	14	US-10-311-455-2112
6	1160.8	35.6	7924	16	US-10-257-166-152
7	1160.6	35.7	2021	9	US-09-764-868-88
8	1000.2	30.6	1054	12	US-10-363-616-74
9	599	18.4	601	9	US-09-817-198A-28
10	599	18.4	601	9	US-09-817-198A-29
11	569	17.5	594	14	US-10-029-386-7830
12	540	16.6	601	9	US-09-817-198A-30
13	534.2	16.4	566	9	US-09-764-868-507
14	516.6	15.9	601	9	US-09-817-198A-31
15	477.4	14.7	481	9	US-09-920-300A-303

16	477.4	14.7	481	13	US-10-033-528-303	Sequence 303, Appl
17	477.4	14.7	481	14	US-10-099-926-303	Sequence 303, Appl
18	438.8	13.5	463	9	US-09-964-824A-26	Sequence 26, Appl
19	424.4	13.0	458	10	US-09-918-995-20848	Sequence 20848, A
20	366.8	9.4	310	9	US-09-878-178-1930	Sequence 1930, Ap
21	366.8	9.4	310	13	US-10-046-935-1930	Sequence 1930, Ap
22	366.8	9.4	310	14	US-10-146-502-1930	Sequence 1930, Ap
23	228.4	7.0	2411	15	US-10-264-049-290	Sequence 290, Appl
24	221.8	6.8	358	12	US-10-240-425-504	Sequence 504, Appl
25	217.8	6.7	624	9	US-09-794-257-9	Sequence 9, Appl
26	217.8	6.7	1161	9	US-09-794-257-7	Sequence 7, Appl
27	217.8	6.7	2497	9	US-09-834-975-879	Sequence 879, Appl
28	217.8	6.7	2497	9	US-09-834-975-885	Sequence 885, Appl
29	217.8	6.7	2497	9	US-09-834-975-894	Sequence 894, Appl
30	217.8	6.7	2497	9	US-09-834-975-896	Sequence 896, Appl
31	210	6.4	601	9	US-09-817-198A-32	Sequence 32, Appl
32	198.4	6.1	1537	9	US-09-925-300-631	Sequence 631, Appl
33	195.6	6.0	3164	14	US-10-096-534-38	Sequence 38, Appl
34	192	5.9	995	14	US-10-084-817-256	Sequence 256, Appl
35	192	5.9	1274	9	US-09-925-302-91	Sequence 91, Appl
36	180	5.5	621	14	US-10-128-714-2241	Sequence 2241, Ap
37	180	5.5	621	14	US-10-128-714-7600	Sequence 7600, Ap
38	174.8	5.4	967	12	US-10-425-114-21224	Sequence 21224, A
39	173.4	5.3	651	9	US-09-938-842A-836	Sequence 836, Appl
40	173.4	5.3	651	11	US-09-938-842A-836	Sequence 836, Appl
41	169.8	5.2	374	12	US-10-085-783A-56703	Sequence 56703, A
42	169.8	5.2	374	15	US-10-242-535A-56703	Sequence 56703, A
43	167	5.1	941	12	US-10-425-114-24435	Sequence 24435, A
44	167	5.1	968	12	US-10-425-114-12529	Sequence 12529, A
45	167	5.1	1162	12	US-10-424-599-58679	Sequence 58679, A

ALIGNMENTS

RESULT 1  
US-09-817-198A-1  
; Sequence 1, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817,198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3257  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-817-198A-1

Query Match	100.0%;	Score 3257;	DB 9;	Length 3257;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3257;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TGCCCGCTGCCCGCCCGCAGTTC	CGCGCCCGCCCGCAGTTCATGCGCCCGCAGTTCATGCGCGAAGCAGTACG	60
Db	1	TGCCCGCTGCCCGCCCGCAGTTC	CGCGCCCGCCCGCAGTTCATGCGCCCGCAGTTCATGCGCGAAGCAGTACG	60
Qy	61	ATGTGCTGTTCCGGCTGCTGCTG	ATGATCGGGGATCGCGGGTGGGGAAGACCTGCTGCTGT	120
Db	61	ATGTGCTGTTCCGGCTGCTGCTG	ATGATCGGGGATCGCGGGTGGGGAAGACCTGCTGCTGT	120
Qy	121	GCCGCTTCCACGACGAGTTCCAC	TCTCGACATCTCCACCATCGGTGTGACTTTA	180
Db	121	GCCGCTTCCACGACGAGTTCCAC	TCTCGACATCTCCACCATCGGTGTGACTTTA	180
Qy	181	AGATGAAGACCATAGAGTAGACG	CGGATCAAAGTCGGATACAGATCTGGGACACTGCAG	240

181 AGATGAAGACCATAGAGGTAGACGGCATCAAGTGGGATACAGATCTGGGACATGCGAG 240 Db  
241 GGCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGCGGCCACAGGGATATTTT 300 Qy  
241 GGCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGCGGCCACAGGGATATTTT 300 Db  
301 TGGTCTATGACATTAGCAGGAGCGCTCTTACCAGCACATCATGAAAGTGGGTGAGTGAAG 360 Qy  
301 TGGTCTATGACATTAGCAGGAGCGCTCTTACCAGCACATCATGAAAGTGGGTGAGTGAAG 360 Db  
361 TGGATGATGACGACGAGAGCGCTCAGAAAGTCTTATTTGGGAATAAGGCTGATGAGG 420 Qy  
361 TGGATGATGACGACGAGAGCGCTCAGAAAGTCTTATTTGGGAATAAGGCTGATGAGG 420 Db  
421 AGCAGAAACGCGAGTGGGAG 480 Qy  
421 AGCAGAAACGCGAGTGGGAG 480 Db  
481 ACTTCTATGAACAAAGTGCCTGCAACCACTCAACATTTAAAGAGTCAATTCAACGCTCTGA 540 Qy  
481 ACTTCTATGAACAAAGTGCCTGCAACCACTCAACATTTAAAGAGTCAATTCAACGCTCTGA 540 Db  
541 CAGAGCTGGTGTCTGAGCGCCCATAGGAAGAGCTGGAAGGCTCCGGATGGTGCAGCA 600 Qy  
541 CAGAGCTGGTGTCTGAGCGCCCATAGGAAGAGCTGGAAGGCTCCGGATGGTGCAGCA 600 Db  
601 ATGAGTTGGCACTGGCAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660 Qy  
601 ATGAGTTGGCACTGGCAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660 Db  
661 CTTTCGAAAACTGT 720 Qy  
661 CTTTCGAAAACTGT 720 Db  
721 CTCAGAGGCGCCGTGGGAGACAGAGGAGCGCGGCTTTGGCCCTGTGTGTGTGTGTGTGT 780 Qy  
721 CTCAGAGGCGCCGTGGGAGACAGAGGAGCGCGGCTTTGGCCCTGTGTGTGTGTGTGTGT 780 Db  
781 GTGATGACCTTATTTAGTATCAGTAGCCACTACCTCCCTGCTGCGCTGGCCCTGAGAGCG 840 Qy  
781 GTGATGACCTTATTTAGTATCAGTAGCCACTACCTCCCTGCTGCGCTGGCCCTGAGAGCG 840 Db  
841 CTGCTGTCTATCAAGCAGCGCTGTCTCCCAAGCGGCTCCACCTTGGAGTGTCTTTCTCA 900 Qy  
841 CTGCTGTCTATCAAGCAGCGCTGTCTCCCAAGCGGCTCCACCTTGGAGTGTCTTTCTCA 900 Db  
901 GCCTGTTTCCCGAGCAGCGCTGTCTACGACCCCGACGATGTGCGCGAAGCACTGTCTC 960 Qy  
901 GCCTGTTTCCCGAGCAGCGCTGTCTACGACCCCGACGATGTGCGCGAAGCACTGTCTC 960 Db  
961 ACCATCCGCAACCCACAGACAGAGCGGCTGGAGTCCAGGCCACTTTTCAGCTGCTC 1020 Qy  
961 ACCATCCGCAACCCACAGACAGAGCGGCTGGAGTCCAGGCCACTTTTCAGCTGCTC 1020 Db  
1021 CTTTCTCGTGCATCGTGTCTTCTGTCTGTCTTCTCTCTTCCGCCCACTTCTCTTCTC 1080 Qy  
1021 CTTTCTCGTGCATCGTGTCTTCTGTCTGTCTTCTCTCTTCCGCCCACTTCTCTTCTC 1080 Db  
1081 TGACCCCTCCCTCCGCTGCGTTTGGTATCAAGCTCTCAAAACCCGCTCCCGGTGTGT 1140 Qy  
1081 TGACCCCTCCCTCCGCTGCGTTTGGTATCAAGCTCTCAAAACCCGCTCCCGGTGTGT 1140 Db  
1141 CTTGCTGTGTGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200 Qy  
1141 CTTGCTGTGTGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200 Db  
1201 GCTCTGGGAGGTTCACCCCTTGGATCCAGGAAGACCTCCACCTCGCTCGTGGGTG 1260 Qy  
1201 GCTCTGGGAGGTTCACCCCTTGGATCCAGGAAGACCTCCACCTCGCTCGTGGGTG 1260 Db  
1261 GGCCAAAGGCTACAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320 Qy  
1261 GGCCAAAGGCTACAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320 Db

1321 TGGGCCTGCTCTCCCACTGACCTTGGAAAAGTGGAGCATCGAGGTAGGAGGAAAACAGCAA 1380 Qy  
1321 TGGGCCTGCTCTCCCACTGACCTTGGAAAAGTGGAGCATCGAGGTAGGAGGAAAACAGCAA 1380 Db  
1381 CCGGGAGTCTCTGAGCCTGGGGCTGCCCTTACCTTACCCATTTCCCGGACCAAGCTTTG 1440 Qy  
1381 CCGGGAGTCTCTGAGCCTGGGGCTGCCCTTACCTTACCCATTTCCCGGACCAAGCTTTG 1440 Db  
1441 CCTTCTTGGCTGCGCGCTGCTCTTTGGGAACTGAGCTCAGAGGCAAGTGTCTTCAG 1500 Qy  
1441 CCTTCTTGGCTGCGCGCTGCTCTTTGGGAACTGAGCTCAGAGGCAAGTGTCTTCAG 1500 Db  
1501 AGAAGGAAACAAATGAGGGGTGGCAGGGATAAAAAGTCACTTCTTCTACCTCCCA 1560 Qy  
1501 AGAAGGAAACAAATGAGGGGTGGCAGGGATAAAAAGTCACTTCTTCTACCTCCCA 1560 Db  
1561 TGAGCATGAACAATTTCTCTCACTGGCTCCCAATTTAAAGATGTGGACCAAGGC 1620 Qy  
1561 TGAGCATGAACAATTTCTCTCACTGGCTCCCAATTTAAAGATGTGGACCAAGGC 1620 Db  
1621 CTGTGGGTACTCTCAGGGGCAAGGAGAGCCTGGGGTCAGTGACACTGTCAAGGCCAACAT 1680 Qy  
1621 CTGTGGGTACTCTCAGGGGCAAGGAGAGCCTGGGGTCAGTGACACTGTCAAGGCCAACAT 1680 Db  
1681 GCACTCCACAAAGGGGAGCAATTTGGAATGGAAGTCACTAGCTCTATGTATCAGGTTAAGA 1740 Qy  
1681 GCACTCCACAAAGGGGAGCAATTTGGAATGGAAGTCACTAGCTCTATGTATCAGGTTAAGA 1740 Db  
1741 GCAAGGAGAGCTGGCCAGGGAAGCAGGTTTGCACAGCAGAGGGGATGTAGCAACAGCA 1800 Qy  
1741 GCAAGGAGAGCTGGCCAGGGAAGCAGGTTTGCACAGCAGAGGGGATGTAGCAACAGCA 1800 Db  
1801 GGGCCTCTAGGCGCCCTCTTCCATTTTAGTGAAGAGAGCATTTCTCAGACTCCCA 1860 Qy  
1801 GGGCCTCTAGGCGCCCTCTTCCATTTTAGTGAAGAGAGCATTTCTCAGACTCCCA 1860 Db  
1861 GGGCCTCTAGGCGCCCTCTTCCATTTTAGTGAAGAGAGCATTTCTCAGACTCCCA 1920 Qy  
1861 GGGCCTCTAGGCGCCCTCTTCCATTTTAGTGAAGAGAGCATTTCTCAGACTCCCA 1920 Db  
1921 GAGAGGGCAAGACTTTCATGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980 Qy  
1921 GAGAGGGCAAGACTTTCATGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980 Db  
1981 CTGTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCTCTTACCTCTCTGATAGCTT 2040 Qy  
1981 CTGTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCTCTTACCTCTCTGATAGCTT 2040 Db  
2041 GAGGTTTGCACCACTGTGGCTACAGGTGGAGGGAAGAGAGACTCTTCTTCTCCAGAG 2100 Qy  
2041 GAGGTTTGCACCACTGTGGCTACAGGTGGAGGGAAGAGAGACTCTTCTTCTCCAGAG 2100 Db  
2101 TGCTATGTTTCAAGAGTTTCTTTAAACCCCATATGGCCCAAGTAGCTCTGAGAGAGGCC 2160 Qy  
2101 TGCTATGTTTCAAGAGTTTCTTTAAACCCCATATGGCCCAAGTAGCTCTGAGAGAGGCC 2160 Db  
2161 TTTTAAAGACGGAACAGTAATTTTACAGTTTCTTCTGGGGTCTTCTGGCCACCGTCCCAAGG 2220 Qy  
2161 TTTTAAAGACGGAACAGTAATTTTACAGTTTCTTCTGGGGTCTTCTGGCCACCGTCCCAAGG 2220 Db  
2221 TGGGCGAGGCTTAGGAAGAGGGTCAATTTTAAAGCCACACATTTAGCTGCAGTGGCTG 2280 Qy  
2221 TGGGCGAGGCTTAGGAAGAGGGTCAATTTTAAAGCCACACATTTAGCTGCAGTGGCTG 2280 Db  
2281 CAGCAAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGGCA 2340 Qy  
2281 CAGCAAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGGCA 2340 Db  
2341 CTCATATGTGAAGATGAAGACCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400 Qy  
2341 CTCATATGTGAAGATGAAGACCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400 Db

```
Qy 2401 ACCATCAAAACCTTTCTCTGACTTACCAAAACAGGAAACAGCAGGAGAGGGTGGCTCA 2460
Db 2401 ACCATCAAAACCTTTCTCTGACTTACCAAAACAGGAAACAGCAGGAGAGGGTGGCTCA 2460
Qy 2461 GGACTTAGGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGAGACGAGAGTTGTAA 2520
Db 2461 GGACTTAGGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGAGACGAGAGTTGTAA 2520
Qy 2521 ATCACTGGCTAATGAGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580
Db 2521 ATCACTGGCTAATGAGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580
Qy 2581 AGTTGCTTCTTGAAGTGGGACTCTTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640
Db 2581 AGTTGCTTCTTGAAGTGGGACTCTTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640
Qy 2641 GCTAGGGAAGTAGGTGATGCCAGGCCCTCAAGCTCTTTCAGCCAGGACTTGAAGATT 2700
Db 2641 GCTAGGGAAGTAGGTGATGCCAGGCCCTCAAGCTCTTTCAGCCAGGACTTGAAGATT 2700
Qy 2701 ATATTGGCAGTGGCTCCAACTCTGGACCAAGTATTTTCAGCTTTCCCTGAAGATCAGGCA 2760
Db 2701 ATATTGGCAGTGGCTCCAACTCTGGACCAAGTATTTTCAGCTTTCCCTGAAGATCAGGCA 2760
Qy 2761 GGGTGCCATTCAATTGTCTTCTCTAGCCCCCTCAGGAAAGAGACTATATTGTATC 2820
Db 2761 GGGTGCCATTCAATTGTCTTCTCTAGCCCCCTCAGGAAAGAGACTATATTGTATC 2820
Qy 2821 TGTACCTTAGGGTCTTGGAAAGGAAACATGGATCAGGATTTCTATAGACTGATAGGCC 2880
Db 2821 TGTACCTTAGGGTCTTGGAAAGGAAACATGGATCAGGATTTCTATAGACTGATAGGCC 2880
Qy 2881 CTATCCCAAGGGCCATGACTGGGAAAGATGATGGAGCAGAGAGAAATTTGGGATTTTA 2940
Db 2881 CTATCCCAAGGGCCATGACTGGGAAAGATGATGGAGCAGAGAGAAATTTGGGATTTTA 2940
Qy 2941 GGGTGCACTACGCTCAACCTTAACTTTTGGTGGCCCTGGGGCATGTCTTTGAGGCCAGAC 3000
Db 2941 GGGTGCACTACGCTCAACCTTAACTTTTGGTGGCCCTGGGGCATGTCTTTGAGGCCAGAC 3000
Qy 3001 TGTAAAGCAGCTCTGTGGCTGTCTTACTCGTACCACTCTGCACTCTGCTGTCTTGAG 3060
Db 3001 TGTAAAGCAGCTCTGTGGCTGTCTTACTCGTACCACTCTGCACTCTGCTGTCTTGAG 3060
Qy 3061 ACTCCATCCAGCCCGCAGCAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGAGG 3120
Db 3061 ACTCCATCCAGCCCGCAGCAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGAGG 3120
Qy 3121 GTGAACCTTCGTGTACTGTGTCTCGGGTCCATATATGAATTTGAGCAGGGTTCTATT 3180
Db 3121 GTGAACCTTCGTGTACTGTGTCTCGGGTCCATATATGAATTTGAGCAGGGTTCTATT 3180
Qy 3181 TTAACACAGATGTTTACAAAATAAAGATTATTTCAACACCAACCAAAAAA 3240
Db 3181 TTAACACAGATGTTTACAAAATAAAGATTATTTCAACACCAACCAAAAAA 3240
Qy 3241 AAAAAAAAAAAAAA 3257
Db 3241 AAAAAAAAAAAAAA 3257
```

RESULT 2

```
US-09-817-198A-3
; Sequence 3, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
```

```
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28770
; TYPE: DNA
; ORGANISM: Human
US-09-817-198A-3
```

```
Query Match 82.8%; Score 2696.6; DB 9; Length 28770;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2717; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 473 TGGCATGGACTTCTATCAAAACAAGTGCCTGCACCAACTCAACATTAAAGATCATTTCC 532
Db 25562 TGCAGGGGAGAGGAGGAGACACTGGACTAACCTGTGCCCTTTGGTTTCCAGTCAATTCAC 25621
Qy 533 GCGTCTCAGCAGAGCTGTGTGTCAGAGGCCCATAGGAAGAGCTGGAAGCCTCCGGATGCG 592
Db 25622 GCGTCTCAGCAGAGCTGTGTGTCAGAGGCCCATAGGAAGAGCTGGAAGCCTCCGGATGCG 25691
Qy 593 TGCAGCAATGAGTTGGCACTTGGCAGAGCTGGAGAGGAGAGGGGCAAAACCCGAGGGCCC 652
Db 25682 TGCAGCAATGAGTTGGCACTTGGCAGAGCTGGAGAGGAGAGGGGCAAAACCCGAGGGCCC 25741
Qy 653 AGCGAACTCTTCGAAAACCTGTGCTGAGTCTGTGTGGGGACACCCACACAGACACC 712
Db 25742 AGCGAACTCTTCGAAAACCTGTGCTGAGTCTGTGTGGGGACACCCACACAGACACC 25801
Qy 713 CCTCTTCCCTCAGAGAGCCCGTGGGACAGACAGGGAGGCGGGGCTTTGGCCCTGTGCTGT 772
Db 25802 CCTCTTCCCTCAGAGAGCCCGTGGGACAGACAGGGAGGCGGGGCTTTGGCCCTGTGCTGT 25861
Qy 773 CCTCTGCTGTGATGATCACTTATTGAGTATCAGTAGCCACTACTCTCCCTGCTGGCCCTGA 832
Db 25862 CCTCTGCTGTGATGATCACTTATTGAGTATCAGTAGCCACTACTCTCCCTGCTGGCCCTGA 25921
Qy 833 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCGTCCAACCTCGAGTGGT 892
Db 25922 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCGTCCAACCTCGAGTGGT 25981
Qy 893 CTCTCTCAGCCCTGTTTCCCGAGCCACAGGCCCTGTACGACCCCGCCACGATGTGCCGCAAGC 952
Db 25982 CTCTCTCAGCCCTGTTTCCCGAGCCACAGGCCCTGTACGACCCCGCCACGATGTGCCGCAAGC 26041
Qy 953 ACTGCTCAACCATCCGACCCACAGACAAACAGCAGGGCTGGAGTCCAGGCCACTTTTC 1012
Db 26042 ACTGCTCAACCATCCGACCCACAGACAAACAGCAGGGCTGGAGTCCAGGCCACTTTTC 26101
Qy 1013 AGCTGCTCTCTTCTCCGTGCACTGTCTTCTCTGCTTTTCTCTCTTCTCTCCCGACATTC 1072
Db 26102 AGCTGCTCTCTTCTCCGTGCACTGTCTTCTCTGCTTTTCTCTCTTCTCTCTCCCGACATTC 26161
Qy 1073 TCTTTCTGACCCCTCCCTCCCGTGGGTTTCTGATCAAAAGCTCTCAAAACCCCGTCCC 1132
Db 26162 TCTTTCTGACCCCTCCCTCCCGTGGGTTTCTGATCAAAAGCTCTCAAAACCCCGTCCC 26221
Qy 1133 CCGTGTGCTCTGTGTGTCAGCTCGCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1192
Db 26222 CCGTGTGCTCTGTGTGTCAGCTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 26281
Qy 1193 GGACCCAGGCTCGTGGGAGGTTCCACCTTTGGATCCAGGAAGAACCTTCCACCTCGCT 1252
Db 26282 GGACCCAGGCTCGTGGGAGGTTCCACCTTTGGATCCAGGAAGAACCTTCCACCTCGCT 26341
Qy 1253 CGTGGGTGGGCAAAAGGCTACAGGGTGTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1312
Db 26342 CGTGGGTGGGCAAAAGGCTACAGGGTGTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 26401
Qy 1313 ATGTGCCATGGCCCTGCTCCCGAGTGCCTGCGAAGTGGAGCATCCAGGTAGGAGGA 1372
Db 26402 ATGTGCCATGGCCCTGCTCCCGAGTGCCTGCGAAGTGGAGCATCCAGGTAGGAGGA 26461
Qy 1373 AACAGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCTATTTCCCGACCA 1432
```

Db 26462 AACGGCAACAGGGAGTCTCGAGCTTGGGGCTGCCCTACCTCTACCCATTCCTCCCGACCA 26521  
Qy 1433 GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCCCTCTCTTTGGGGAAGTGAAGCTCAGAGGCGAG 1492  
Db 26522 GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCCCTCTCTTTGGGGAAGTGAAGCTCAGAGGCGAG 26581  
Qy 1493 TGCTTCAGAGAAGGAACAAATAGAGGGTGGCAGGGATAAAGTCAACCTCCATCTCT 1552  
Db 26582 TGCTTCAGAGAAGGAACAAATAGAGGGTGGCAGGGATAAAGTCAACCTCCATCTCT 26641  
Qy 1553 ACCTCCCATGAGATGAACAAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGG 1612  
Db 26642 ACCTCCCATGAGATGAACAAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGG 26701  
Qy 1613 ACCAAGGCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGTGAGTGAACCTGTCAAG 1672  
Db 26702 ACCAAGGCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGTGAGTGAACCTGTCAAG 26761  
Qy 1673 CCAACCATGCACTCCAAAGGGGAGCATTTGGAAATGAAGGACTAGCTCTCTATGTATCA 1732  
Db 26762 CCAACCATGCACTCCAAAGGGGAGCATTTGGAAATGAAGGACTAGCTCTCTATGTATCA 26821  
Qy 1733 GGTTAAGACAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGAGAGGGGAATGTAG 1792  
Db 26822 GGTTAAGACAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGAGAGGGGAATGTAG 26881  
Qy 1793 CAACAGCAGGGCTCTAGAGCCCATCTTCCATTTCTTAGTGAAGAGCATTTCCCTCA 1852  
Db 26882 CAACAGCAGGGCTCTAGAGCCCATCTTCCATTTCTTAGTGAAGAGCATTTCCCTCA 26941  
Qy 1853 GACTCCAGGCGAGGACTAGAGCTAGCTTACCAACCAAGGTTCTCTGGGACCCAAA 1912  
Db 26942 GACTCCAGGCGAGGACTAGAGCTAGCTTACCAACCAAGGTTCTCTGGGACCCAAA 27001  
Qy 1913 GTTTATGGGAGAAGGCAAGACTTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 1972  
Db 27002 GTTTATGGGAGAAGGCAAGACTTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 27061  
Qy 1973 GCTTGTGTCTGTCTCTTGGCTTTAAGACAAGCGCTCATCTTGGCTCTACTCTCTG 2032  
Db 27062 GCTTGTGTCTGTCTCTTGGCTTTAAGACAAGCGCTCATCTTGGCTCTACTCTCTG 27121  
Qy 2033 ATAGCTTGAGGGTTTGCACACACACTGTGGCTACAGAGTGGAGGAGGAGGAGGAGGAGG 2092  
Db 27122 ATAGCTTGAGGGTTTGCACACACACTGTGGCTACAGAGTGGAGGAGGAGGAGGAGGAGG 27181  
Qy 2093 CTCAGAGTGTATGTTCAGGAAGTTTCTTAAACCCCATATGGCCCAAGAGTGTCTCGTA 2152  
Db 27182 CTCAGAGTGTATGTTCAGGAAGTTTCTTAAACCCCATATGGCCCAAGAGTGTCTCGTA 27241  
Qy 2153 GGAGGCCCTTTAAGACGGACAGTAATTTACAGTTCTACTGGGGTCTCTGCCACCG 2212  
Db 27242 GGAGGCCCTTTAAGACGGACAGTAATTTACAGTTCTACTGGGGTCTCTGCCACCG 27301  
Qy 2213 TCCCAAGTGGCGAGGCTAGGAGAGGGTCACTTTAAGCCACACATAGCTGCACTG 2272  
Db 27302 TCCCAAGTGGCGAGGCTAGGAGAGGGTCACTTTAAGCCACACATAGCTGCACTG 27361  
Qy 2273 CGTGGCTGCAGCCAAAACAAAGAACTGGGTGTAGTATTCACTAAGAACCAAAAT 2332  
Db 27362 CGTGGCTGCAGCCAAAACAAAGAACTGGGTGTAGTATTCACTAAGAACCAAAAT 27421  
Qy 2333 CCAGGCACTATATGTGAAGATGAAGACTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2392  
Db 27422 CCAGGCACTATATGTGAAGATGAAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 27481  
Qy 2393 GGGAAAGAACCATCAAACTTCT 2452  
Db 27482 GGGAAAGAACCATCAAACTTCT 27541  
Qy 2453 GTGGCTCAGGACTTAGGACAGGGTATAGTGTAGTGTGAAGAGCAAGAGGAGGAGGAGGAGGAGG 2512

Db 27542 GTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGAGCAGGA 27601  
Qy 2513 AGTTGTAATCACTGGCTAATGAGAAAGAGAGCAGCTAACTCTAGGATGAAGCTGTGAC 2572  
Db 27602 AGTTGTAATCACTGGCTAATGAGAAAGAGAGCAGCTAACTCTAGGATGAAGCTGTGAC 27661  
Qy 2573 TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632  
Db 27662 TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 27721  
Qy 2633 ACACCTGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTT 2692  
Db 27722 ACACCTGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTT 27781  
Qy 2693 GAGAAGTTATTTGGGCGAGTGGCTCCAATCTGTGGACCAAGTATTTTCAAGCTTTCCCTCAAG 2752  
Db 27782 GAGAAGTTATTTGGGCGAGTGGCTCCAATCTGTGGACCAAGTATTTTCAAGCTTTCCCTCAAG 27841  
Qy 2753 ATCAGGAGGGTGCCTATTTCTTCTCTCTAGCCCCCTCAGGAAAGAGGACTAT 2812  
Db 27842 ATCAGGAGGGTGCCTATTTCTTCTCTCTAGCCCCCTCAGGAAAGAGGACTAT 27901  
Qy 2813 ATTTGTACTGTACCTAGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACT 2872  
Db 27902 ATTTGTACTGTACCTAGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACT 27961  
Qy 2873 GATAGGCCCTTATCCAAAGGGCCATGACTGGGAAAGGTATGGAGCAGAGGAGAAATTG 2932  
Db 27962 GATAGGCCCTTATCCAAAGGGCCATGACTGGGAAAGGTATGGAGCAGAGGAGAAATTG 28021  
Qy 2933 GGAATTTAGGGTGCAGTACGCTCACCTTAACTTTTGTGGCTGGGGCATGTCTTGAG 2992  
Db 28022 GGAATTTAGGGTGCAGTACGCTCACCTTAACTTTTGTGGCTGGGGCATGTCTTGAG 28081  
Qy 2993 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCCACCCTCTGCACCTGCT 3052  
Db 28082 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCCACCCTCTGCACCTGCT 28141  
Qy 3053 GTCTGAGACTCCATCCAGCCCCCAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3112  
Db 28142 GTCTGAGACTCCATCCAGCCCCCAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 28201  
Qy 3113 TGTGAGGGTGAACCTTGTGTACTGTGTCTCGGGTCCATATATGAATTTGAGCAGGGTT 3172  
Db 28202 TGTGAGGGTGAACCTTGTGTACTGTGTCTCGGGTCCATATATGAATTTGAGCAGGGTT 28261  
Qy 3173 CATCTATTTTAAACACAGATGTTTACAAATGAAGTATTTTCAAAACCC 3223  
Db 28262 CATCTATTTTAAACACAGATGTTTACAAATGAAGTATTTTCAAAACCC 28312

## RESULT 3

US-10-311-455-2111  
; Sequence 2111, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Cytosine  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311.455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 2111  
; LENGTH: 7924

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2111

Query Match          36.6%; Score 1192.8; DB 14; Length 7924;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 1476; Conservative 0; Mismatches 472; Indels 0; Gaps 0;

QY 1276 GGTGCTCTTCCCTCTCCCCACCCCCACATGTCCTCATGTGCGCATGGGCTGCTCCCC 1335
    |||||
Db 1 GGTGTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTTATTTTATGTTGTTATGGGTTGTTTTT 60

QY 1336 AGTGAOCTGCGAAGTGGAGCATCGAGTAGGAGGGAACAGCAACCGGGAGTCTCTCGA 1395
    |||||
Db 61 AGTGATTTGCGAAGTGGAGTATCGAGGTAGGAGGGAACCGTAAATTTAGGAGTTCGGA 120

QY 1396 GCCTGGGCTGCCCTACCTCTACCCATTCGCCGACGAGCTTTGCCCTTGCTTGCTGCTC 1455
    |||||
Db 121 GTTTGGGGTTGTTTTTATTTTATTTTATTTTTCGATTAGAGTTTGTGTTTTGTTGGTTGT 180

QY 1456 CCSCCTGCCCTTTGGGGAACCTGAGCTCAGAGCAGGTGCTTCAGAGAAGGAACAAAT 1515
    |||||
Db 181 TCGTTTTGTTTTTTGGGGAATGAGTTTGAAGGTAGGTGTTTTTAGAGAAGAATAAAAT 240

QY 1516 GAGGGGTGGCAGGATAAAAAGTCACCTCAATTCCTCACTCCCATCGCATGAAACACA 1575
    |||||
Db 241 GAGGGGTGGTAGGATAAAAAGTTATTTTATTTTATTTTATGTTAGTAGTATGAATATA 300

QY 1576 ATTTCTCTCACCTGGCTCCCAATTTAAAGATGTGGACCAAGCCCTGTGGGTACTCCAG 1635
    |||||
Db 301 ATTTTTTTTTTATTTGGTTTTTAAATTTAAAGATGTGGAATTAAGGTTTGTGGGTATTTAG 360

QY 1636 GGGCAGGAGAGCCCTGGGCTCAGTGACCTGTCCAGCCACCACTCCCAAGGG 1695
    |||||
Db 361 GGGTAAGGAGAGTTTGGGGTTAGTGATATTTGTTAGGTTAATTATGTTATTTTATAAGGG 420

QY 1696 GAGCATTTGGAATTAAGAGCTAGCTCCTATGTATCAGGTTAAAGACAAGGAGAGCTGG 1755
    |||||
Db 421 GAGTATTTGGAATGAAGATAGTTTTTATGTATTTAGTTTAAAGTTAAGGAGAGTTGG 480

QY 1756 CCAAGGACAGAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGGCCTCTTAGGCC 1815
    |||||
Db 481 TTAGGGATAGTAGTTTGTATAGTAGAGGGGAATGTAGTAATAGTAGGGTTTTTTAGGTTT 540

QY 1816 CATCTTCATTTCTTAGTGAAGAGACATTTCTCTCAGATCCCGAGCGGAGGACTGAGC 1875
    |||||
Db 541 TATTTTTTATTTTTTAGTTAAGAAGAGTATTTTTTTTTAGTTAGTTTGTAGCGGAGGATTGAGT 600

QY 1876 CTAGCCTCTCAGCAACCAAGTTCTCTCGGACCCAAAGTTTTATGGGAGAAGGGCAAGAC 1935
    |||||
Db 601 TTAGTTTTTAGTAAATTAAGTTTTTTTTTGGATTTAAAGTTTTATGGGAAGAGGTTAAGAT 660

QY 1936 TTCATGGGAAGAGAGAAGGAGCCCTGGGTAGAAACGCTTGGTGCTGTTCTCTTTGGCC 1995
    |||||
Db 661 TTTATGGGAAGAGAGAAGGAAGGTTTTTGGGTAGAAACGTTTTGGTGTGTTTTTTTTTGGTT 720

QY 1996 TTTAAGCAAAAGCGCTCATCTGCGCTCTACCTCCTGATAGGCTTGAGGGTTTGCACAAC 2055
    |||||
Db 721 TTTAAGATAAAGCGTTTATTTTGTTTTTTATTTTTTGTATAGTTTGTAGGGTTTGTTAATT 780

QY 2056 ACACTGTGGCTACAGGTGGAGGGAAGAGGACTCCTCTCCAGAGTGCTATGTTTCAGAA 2115
    |||||
Db 781 ATATTGGGTTTATAGGTGGAGGGAAGAGGATTTTTTTTTTTAGAGTGTATGTTTAGGAA 840

QY 2116 GTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCCCCCTTAAAGACGGAACA 2175
    |||||
Db 841 GTTTTTTTTAAATTTATATGTTTAAAGAGTAGTTCGTAGGAGGTTTTTTTTTAAAGACGGAATA 900

QY 2176 AGTAAATTTACAGTTCTACTTGGGGTTCTCTGCCACCGTCCCAAGTGGGCGAGGCCTAGG 2235
    |||||
Db 901 AGTAAATTTATAGTTTATTTAGGTTTGTGTTTATCGTTTAAAGTGGGCGAGGTTTAGG 960

```

RESULT 4  
US-10-257-166-151  
; Sequence 151, Application US/10257166





Db	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779	1780
QY	AGGCAGCCACCTGCTCCTGAGCCTCCACATATCTCCTGTGACGGGTGAACCTTCGTGTAC	3135																																						
Db	1801	AGGTAGTATTTGTTTGGTATTTATTTATTTTGGTATTTATTTAGTATTTATTTAGTATTT	1860																																					
QY	3136	TGTGCTTCGGTCCATATATGAATGTGACAGGGTTCATCTATTTTAAACACAGATGTT	3195																																					
Db	1861	TGTGTTTCGGGTTATATATGAATGTGAGTAGGGTTATTTATTTTAAATATAGATGTT	1920																																					
QY	3196	TACAAATAAAGATTATTTCAAAACCACC	3223																																					
Db	1921	TATAAAATAAGATTATTTTAAATATTC	1948																																					

RESULT 5  
 US-10-311-455-2112/c  
 ; Sequence 2112, Application US/10311455  
 ; Publication No. US20030143606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIEPENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Cytosine  
 ; TITLE OF INVENTION: cytosine methylation  
 ; FILE REFERENCE: 5013.1014  
 ; CURRENT APPLICATION NUMBER: US/10/311.455  
 ; CURRENT FILING DATE: 2002-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
 ; PRIOR FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: DE 10032529.7  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: DE 10043826.1  
 ; PRIOR FILING DATE: 2000-09-01  
 ; NUMBER OF SEQ.ID NOS: 2424  
 ; SEQ ID NO 2112  
 ; LENGTH: 7924  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-311-455-2112

Query Match	35.6%	Score 1160.8	DB 14	Length 7924
Best Local Similarity	74.1%	Pred. No. 2.4e-309		
Matches 1468	Conservative 0	Mismatches 512	Indels 0	Gaps 0

QY	1278	TGCTTCTTCCTTTCCCCACCCACCTGTCCTCATGTGTCATATGCGCCATGGCGCTGCTCCCCAG	1337
Db	7922	TACTTCTTCCTTCTCCCCACCCACCTATCCTCATATACCATATAAACCTTACCTCCCCAA	7863
QY	1338	TGACCTTCGGAATGGAGCATCGAGGTAGAGGGGAACAGACCGGGGAGTCTCGAGC	1397
Db	7862	TAACTTCAGAAAAATAAAACATCGAAATAAAAAAAGCAACCAAAAAATCCTCGAAC	7803
QY	1398	CTGGGGTGCCCTTACCTTACCCATTCCCGACAGAGCTTTTGGCTTGCTTGGCTGCCC	1457
Db	7802	CTAAAACTACCTTACCTTACCCATTCCCGACCAAACTTTTACCTTACTTAACTACCC	7743
QY	1458	GCTGCTCTTTTGGGGAATGAGCTCAGAGGCGAGGTCTTTCAGAGAAGGAAACAAATGA	1517
Db	7742	GCCTACTCTTTTAAAAAACTTAACTCAAAACAAATACTTTCAAAAAAATAA	7683
QY	1518	GGGGTGGCAGGATATAAAGTCACTTCTTCTTACCTCCCATGAGCATGAACACAT	1577
Db	7682	AAAAATAACAAAAATAAAAAATCACTTCTTCTTACCTCCCATGAGCATGAACACAT	7623
QY	1578	TTCTCTCCACTGGCTCCCAATTTTAAAGATGTGGACCAAGGCTGTGGGTACTCCAGGG	1637
Db	7622	TTCTCTCCACTTAACTCCCAATTTTAAAGATGTGGACCAAGGCTGTGGGTACTCCAGGG	7563
QY	1638	GCAAGGAGAGCCCTGGGGTCAGTGACACTGTCTAGGCGCAACCATGCACTCCCAAAAGGGA	1697









QY 643 CCGAGGCCCGAGCGAACTCTTGAACACCTGCTGGTCTGAGTCTCTGTGGGGCACCCC 702  
Db 146 CCGAGGCCCGAGCGAACTCTTGAACACCTGCTGGTCTGAGTCTCTGTGGGGCACCCC 205  
QY 703 ACACGACACCCCTCTTCCCTCAGAGGCCCGTGGGCAGACAGGGAGCCCGGGGCTTTGCC 762  
Db 206 ACACGACACCCCTCTTCCCTCAGAGGCCCGTGGGCAGACAGGGAGCCCGGGGCTTTGCC 265  
QY 763 CTGCTGCTGCTCTCTGCTGATGACCTATTAGTATCAGTAGCCACTACTCCCTCTGC 822  
Db 266 CTGCTGCTGCTCTCTGCTGATGACCTATTAGTATCAGTAGCCACTACTCCCTCTGC 325  
QY 823 CTGSCCTCAGAGGGGCTGCTGCTCATCTCAAGCAGCCCTGTCGCCAGCCGCTCCACC 882  
Db 326 CTGSCCTCAGAGGGGCTGCTGCTCATCTCAAGCAGCCCTGTCGCCAGCCGCTCCACC 385  
QY 883 CTGAGTGGTCTTCTTCAGGCTGTTTCCCGCAGCCAGGCTGCTACGACCCCGCAGATG 942  
Db 386 CTGAGTGGTCTTCTTCAGGCTGTTTCCCGCAGCCAGGCTGCTACGACCCCGCAGATG 445  
QY 943 TGGCGAAGCACTGCTCTACATCCCGACCCGACACAGCCAGCCAGGGCTGGAGTCCA 1002  
Db 446 TGGCGAAGCACTGCTCTACATCCCGACCCGACACAGCCAGCCAGGGCTGGAGTCCA 505  
QY 1003 GGGCACCTTTCAGTCTCTCTTCTCTGACCCCTCCC 1091  
Db 506 GGGCACCTTTCAGTCTCTCTTCTCTGATCGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1062  
QY 1063 CCCCCACTTCTCTTCTCTGACCCCTCCC 1091  
Db 566 CCCCCACTTCTCTTCTCTGACCCCTCCC 594

## RESULT 12

US-09-817-198A-30  
; Sequence 30, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817,198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-817-198A-30

Query Match 16.6%; Score 540; DB 9; Length 601;  
Best Local Similarity 99.6%; Pred. No. 2e-138;  
Matches 540; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2682 GCCAGGGACTTGAGAAGTATATTGGGAGTGGCTCCAACTGTGGACCAATTTTCAGC 2741  
Db 1 GCCAGGGACTTGAGAAGTATATTGGGAGTGGCTCCAACTGTGGACCAATTTTCAGC 60  
QY 2742 TTTCCCTGAAGATCAGGCGAGGTCCTTTCATTTCTCTCTAGCCCCCTCAGAA 2801  
Db 61 TTTCCCTGAAGATCAGGCGAGGTCCTTTCATTTCTCTCTAGCCCCCTCAGAA 120  
QY 2802 AGAAGGACTATATTGTACTGTACCTAGGGGTTCTGGAAGGGAAAAACATGAATCAGGA 2861  
Db 121 AGAAGGACTATATTGTACTGTACCTAGGGGTTCTGGAAGGGAAAAACATGAATCAGGA 180  
QY 2862 TTCTATAGCTATAGGCGCTTATCCAGAGGCCATGCTGGGAAAGGATGGAGCAG 2921  
Db 181 TTCTATAGCTATAGGCGCTTATCCAGAGGCCATGCTGGGAAAGGATGGAGCAG 240

QY 2922 AAGGAGAAATGGGATTTTAGGGTGCAGCTACGCTACCCCTAAACTTTTGTGGCTGGGG 2981  
Db 241 AAGGAGAAATGGGATTTTAGGGTGCAGCTACGCTACCCCTAAACTTTTGTGGCTGGGG 300  
QY 2982 CATGTCTTGAGGCCAGAGACTGTTAAGCAGGCTCTGTGGCTGTTTACTGCTCACCACCT 3041  
Db 301 YATGTCTTGAGGCCAGAGACTGTTAAGCAGGCTCTGTGGCTGTTTACTGCTCACCACCT 360  
QY 3042 CTGACACTGCTGCTGTGAGACTCCATCCAGCCCGCAGCCAGCCCTGCTCCTGAGCCTC 3101  
Db 361 CTGACACTGCTGCTGTGAGACTCCATCCAGCCCGCAGCCAGCCCTGCTCCTGAGCCTC 420  
QY 3102 CACTATCTCCTCTGACGGGTGAACCTTCGTGTACTGTCTCGGGTCCATATATGAATTG 3161  
Db 421 CACTATCTCCTCTGACGGGTGAACCTTCGTGTACTGTCTCGGGTCCATATATGAATTG 480  
QY 3162 TGAGCAGGGTTCATCTATTTTAAACAAGATGTTTACAAATAAAGATTTATTTCAAACCA 3221  
Db 481 TGAGCAGGGTTCATCTATTTTAAACAAGATGTTTACAAATAAAGATTTATTTCAAACCA 540  
QY 3222 CC 3223  
Db 541 CC 542

## RESULT 13

US-09-764-868-507  
; Sequence 507, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 507  
; LENGTH: 566  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (484)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (493)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (538)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (563)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-868-507

Query Match 16.4%; Score 534.2; DB 9; Length 566;  
Best Local Similarity 98.9%; Pred. No. 7.9e-137;  
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCCCGCTGCCCGCCGCGAGTTCGCGGCCCGCTGGGCCCGCAGTCATGCGGAGCAGTACGA 61  
Db 28 GCCCGCTGCCCGCCGCGAGTTCGCGGCCCGCTGGGCCCGCAGTCATGCGGAGCAGTACGA 87  
QY 62 TGTGCTGTTCCGGCTGCTGCTGATTCGGGACTCTCCGGGGTGGGCAAGACCTGCTGCTGTG 121  
Db 88 TGTGCTGTTCCGGCTGCTGCTGATTCGGGACTCTCCGGGGTGGGCAAGACCTGCTGCTGTG 147  
QY 122 CCGCTTCACCGAACACAGATTCACCTCTCGCACATCTCCACCATCTCGTGTGACTTTAA 181  
Db 148 CCGCTTCACCGAACACAGATTCACCTCTCGCACATCTCCACCATCTCGTGTGACTTTAA 207



QY 182 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 241  
DB 208 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 267  
QY 242 GCAGGAGATACCAAGACCATCAAAAGCAGTACTATCGGGGGCCGACGGGATATTTT 301  
DB 268 GCAGGAGATACCAAGACCATCAAAAGCAGTACTATCGGGGGCCGACGGGATATTTT 327  
QY 302 GGTCTATGACATTAGCAGGAGCGCTTTACAGCACATCATGAAGTGGGTCAAGTACGT 361  
DB 328 GGTCTATGACATTAGCAGGAGCGCTTTACAGCACATCATGAAGTGGGTCAAGTACGT 387  
QY 362 GGATGAGTACGCACAGAGGCGTCCAGAGATCCTTATTGGGAATAAGGCTGATGAGGA 421  
DB 388 GGATGAGTACGCACAGAGGCGTCCAGAGATCCTTATTGGGAATAAGGCTGATGAGGA 447  
QY 422 GCAGAACGGCAGGTGGGAAAGAGAGCAGGAGCGAGCTGGGAGGAGTATGGCATGGA 481  
DB 448 GCAGAACGGCAGGTGGGAAAGAGAGCAGGAGCGAGCTGGGAGGAGTATGGCATGGA 507  
QY 482 CTTCTATGAACAAGTGGTCCACCACTCAACATTAAAGATCATTTACGCGCTCTGA 540  
DB 508 CTTCTATGAACAAGTGGTCCACCACTCAACATTAAAGATCATTTACGCGCTCTGA 566

## RESULT 14

US-09-817-198A-31

; Sequence 31, Application US/09817198A

; Patent No. US20020146758A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL001188

; CURRENT APPLICATION NUMBER: US/09/817,198A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 31

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-817-198A-31

Query Match 15.9%; Score 516.6; DB 9; Length 601;

Best Local Similarity 99.8%; Pred. No. 6e-132;

Matches 516; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2707 GGCAGTGGCTCCAATCTGTGACACAGTATTTTCAGCTTTCCCTGAAGATCAGGCGGGTGC 2766

DB 1 GGCAGTGGCTCCAATCTGTGACACAGTATTTTCAGCTTTCCCTGAAGATCAGGCGGGTGC 60

QY 2767 CATTCAATGCTTTCTCTCTAGCCCTCAGGAAAGAGCTATATTTGTACTGTACC 2826

DB 61 CATTCAATGCTTTCTCTCTAGCCCTCAGGAAAGAGCTATATTTGTACTGTACC 120

QY 2827 CTAGGGGTTCTGGAAGGAAACATGGAATCAGATTTCTATAGACTGATAGCCCTATCC 2886

DB 121 CTAGGGGTTCTGGAAGGAAACATGGAATCAGATTTCTATAGACTGATAGCCCTATCC 180

QY 2887 ACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGAAATTTGGATTTAGGGTGC 2946

DB 181 ACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGAAATTTAGGGTGC 240

QY 2947 AGCTACGCTACCCCTAAACTTTTGGTGGCCCTGAGGAGCTGCTTTCAGGCCCCAGACTGTAA 3006

DB 241 AGCTACGCTACCCCTAAACTTTTGGTGGCCCTGAGGAGCTGCTTTCAGGCCCCAGACTGTAA 300

QY 3007 GCAGGCTCTGTGGCTGTGTACTGTGTCACCACTCTGCACCTGTCTTGTAGACTCCA 3066

DB 301 SCAGGCTCTGTGGCTGTGTACTGTGTCACCACTCTGCACCTGTCTTGTAGACTCCA 360

QY 3067 TCCAGCCCCAGGACGACCACTGCTCTGAGCCTTCCACTATCTCCCTGTGACGGGTGAAC 3126  
DB 361 TCCAGCCCCAGGACGACCACTGCTCTGAGCCTTCCACTATCTCCCTGTGACGGGTGAAC 420  
QY 3127 TTGCTGACTGTGTCTCGGTCCATATATGAATTTGTGAGAGGGTTCATCTATTTTAAAC 3186  
DB 421 TTGCTGACTGTGTCTCGGTCCATATATGAATTTGTGAGAGGGTTCATCTATTTTAAAC 480  
QY 3187 ACAGATGTTTACAAAATAAAGATTTATTTCAAAACCACC 3223  
DB 481 ACAGATGTTTACAAAATAAAGATTTATTTCAAAACCACC 517

## RESULT 15

US-09-920-300A-303

; Sequence 303, Application US/09920300A

; Patent No. US20020136728A1

; GENERAL INFORMATION:

; APPLICANT: King, Gordon E.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Xu, Jiangchun

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.547

; CURRENT APPLICATION NUMBER: US/09/920,300A

; CURRENT FILING DATE: 2001-07-31

; NUMBER OF SEQ ID NOS: 1789

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 303

; LENGTH: 481

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 417, 461

; OTHER INFORMATION: n = A,T,C or G

US-09-920-300A-303

Query Match

14.7%; Score 477.4; DB 9; Length 481;

Best Local Similarity 99.4%; Pred. No. 3.6e-121;

Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACAACCTGGG 2641

DB 1 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACAACCTGGG 60

QY 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGGACTTGAAGTTA 2701

DB 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGGACTTGAAGTTA 120

QY 2702 TATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCCAGCTTTCCCTGAAGATCAGGAG 2761

DB 121 TATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCCAGCTTTCCCTGAAGATCAGGAG 180

QY 2762 GGTGCCATTCAATGCTCTTCTCTCCTAGCCCCCTCAGGAAAGAGGACTATATTTGTACT 2821

DB 181 GGTGCCATTCAATGCTCTTCTCTCCTAGCCCCCTCAGGAAAGAGGACTATATTTGTACT 240

QY 2822 GTACCCCTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2881

DB 241 GTACCCCTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 300

QY 2882 TATCCACAGGGCCATGACTGGGAAAAGGTATGGGACAGAGGAAATTTGGGATTTTAG 2941

DB 301 TATCCACAGGGCCATGACTGGGAAAAGGTATGGGACAGAGGAAATTTGGGATTTTAG 360

QY 2942 GGTGCAGCTACGCTCACCCCTAAACTTTTGTGGCTGGGCGATCTTTGAGGCCCCAGACT 3001

DB 361 GGTGCAGCTACGCTCACCCCTAAACTTTTGTGGCTGGGCGATCTTTGAGGCCCCAGACT 420

QY 3002 GTTAAAGCAGGCTCTGTCTGGGCTGTTTACTCGTCACCACTCTGCACCTGTCTTGAGA 3061

Db 421 GTTAAACAGGCTGCTGGCCCTGTTTACTGCTGCACCACTGTCCTGCTGTCTTGAGA 480

Qy 3062 C 3062

Db 481 C 481

Search completed: March 21, 2004, 05:17:48  
Job time : 1091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 04:56:29 ; Search time 7314 Seconds  
(without alignments)  
13297.938 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcgcgtgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2874.2	88.2	3151	11	CNSLTI18T BX248046 human ful
2	2690.4	82.6	2700	11	BC014511 Homo sapi
3	947.8	29.1	1081	12	BM551913 AGENCOURT
4	883.4	27.1	956	13	BX397021 BX397021

5	856	26.3	957	13	BQ937298	BQ937298 AGENCOURT
6	847.2	26.0	915	13	BX397020	BX397020 BX397020
7	840.2	25.8	908	13	BQ716774	BQ716774 AGENCOURT
8	837.8	25.7	918	13	BX326279	BX326279 BX326279
9	810.6	24.9	901	13	BQ220195	BQ220195 AGENCOURT
10	807.8	24.8	929	13	BX348287	BX348287 BX348287
11	801.6	24.6	924	12	B1118034	B1118034 602867301
12	799.2	24.5	1046	13	BQ277739	BQ277739 AGENCOURT
13	797.2	24.5	840	13	BUS57215	BUS57215 AGENCOURT
14	792.2	24.3	820	13	BUS49866	BUS49866 AGENCOURT
15	792.2	24.3	863	13	BUS57274	BUS57274 AGENCOURT
16	781.2	24.0	1027	12	BM476862	BM476862 AGENCOURT
17	779.6	23.9	907	13	BUS41549	BUS41549 AGENCOURT
18	769.4	23.6	922	13	BUS44769	BUS44769 AGENCOURT
19	767.4	23.6	878	12	BG676934	BG676934 602623549
20	758.4	23.3	928	13	BUI51020	BUI51020 AGENCOURT
21	755.4	23.2	767	12	BM977746	BM977746 UI-CF-EN1
22	746.2	22.9	763	12	BM975016	BM975016 UI-CF-EC1
23	746	22.9	990	10	BE799964	BE799964 601587861
24	743.6	22.8	793	14	CA411919	CA411919 UI-H-E20-
25	742	22.8	756	13	BX093757	BX093757 BX093757
26	733.8	22.5	738	13	BUI72073	BUI72073 AGENCOURT
27	723.6	22.2	733	13	BX331206	BX331206 BX331206
28	722	22.2	732	12	BM045331	BM045331 603623321
29	721.8	22.2	749	12	BG769088	BG769088 602743340
30	716.8	22.0	766	12	BM049329	BM049329 603626279
31	704.6	21.6	729	13	BX352668	BX352668 BX352668
32	704.6	21.6	907	13	BX388608	BX388608 BX388608
33	698.2	21.4	1116	13	BQ067682	BQ067682 AGENCOURT
34	697	21.4	744	12	BG697415	BG697415 602661196
35	695.2	21.3	703	13	BUS67540	BUS67540 UI-CF-EC1
36	689.8	21.2	845	12	BG821168	BG821168 602724570
37	689.4	21.2	783	12	BG760915	BG760915 602717606
38	687.2	21.1	887	12	BG766350	BG766350 602739059
39	687.2	21.1	935	13	BQ876547	BQ876547 AGENCOURT
40	664	20.4	825	14	CD643455	CD643455 AGENCOURT
41	662	20.3	899	13	BUS52993	BUS52993 AGENCOURT
42	644.2	19.8	891	10	BE883791	BE883791 601506187
43	638	19.6	639	13	BQ188717	BQ188717 UI-E-EJ1-
44	632.2	19.4	637	14	CA424721	CA424721 UI-H-FE1-
45	618.2	19.0	647	10	BF966846	BF966846 602286692

ALIGNMENTS

RESULT 1	CNSLTI18T	3151 bp	mRNA	linear	HTC 18-JUN-2003
LOCUS	human full-length cDNA clone CS0DI028YM11 of Placenta of Homo sapiens (human).				
DEFINITION	human full-length cDNA clone CS0DI028YM11 of Placenta of Homo sapiens (human).				
ACCESSION	BX248046				
VERSION	BX248046.1	GI:28193221			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 3151) Genoscope.				
REFERENCE	Submitted (30-JAN-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)				
TITLE	- Web : www.genoscope.cns.fr				
JOURNAL	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned				
COMMENT					

into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

## FEATURES

Location/Qualifiers

source

1..3151  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS01028YM11"  
/tissue\_type="Placenta"  
/notes="Cot 25-normalized-vector pCMVSPORT\_6"  
<1..504  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD62353.1"  
/db\_xref="GI:28193222"  
/translation="GVDFPKMTIEVDGIKVRIQIWDTAGQERYQTITKQYVRAQGLF  
LVYDISERSYQHIMKWSVDVEVDGATSLPGCGEGASPGKARRGPDGKANASRKUL  
PQPMKTSYTHQKASRRSLGILMRSGRWEESKGSWRRSMAMTSMKQVPAPTST  
LKSHSRV"

## CDS

## ORIGIN

Query Match 88.2%; Score 2874.2; DB 11; Length 3151;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 3017; Conservative 0; Mismatches 3; Indels 131; Gaps 1;

Qy	168	GGTGTTCACATTTAAGATGAAGACCATAGAGGTAGACGCGCATCAAGTGCAGATACAGATC	227
Db	1	GGTGTTCACATTTAAGATGAAGACCATAGAGGTAGACGCGCATCAAGTGCAGATACAGATC	60
Qy	228	TGGGACACTGCAGGCGAGGAGATACCAGACCATCAAAAGCAGTACTATCGCGGGGCC	287
Db	61	TGGGACACTGCAGGCGAGGAGATACCAGACCATCAAAAGCAGTACTATCGCGGGGCC	120
Qy	288	CAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTTTACAGGACATCATGAAG	347
Db	121	CAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTTTACAGGACATCATGAAG	180
Qy	348	TGGGTCAGTGACGTGGATGA-----	367
Db	181	TGGGTCAGTGACGTGGATGAGTAGGATAGCCACCTCACTGCCGGGGTGTGGAGAGGT	240
Qy	368	-----	367
Db	241	GCCTCACCGGGGAAGCAAGCGAGGGCCAGATGGGAAGGCAAAATGCTTCCAGGAAGCTT	300
Qy	368	-----GTAGGCACCAAGAGCGGTCCAGAGATCC	396
Db	301	TGCCTTCCACAGCCCTGGATGAAGACCTCTGGTACGCACCAAGAGCGTCCAGAGATCC	360
Qy	397	TTATTGGGAATAAGGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGAGCAAGGCGAGC	456
Db	361	TTATTGGGAATAAGGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGAGCAAGGCGAGC	420
Qy	457	AGCTTGGCGAAGGAGTATGGCATGGAATTTCTATGAAACAAAGTGCCTGCACCAACCTCAACA	516
Db	421	AGCTTGGCGAAGGAGTATGGCATGGAATTTCTATGAAACAAAGTGCCTGCACCAACCTCAACA	480
Qy	517	TTAAAGAGTCAATTCAAGGCTGATGAGAGTGGTGTGCTGCGAGGCCCATAGGAAGAGCTGG	576
Db	481	TTAAAGAGTCAATTCAAGGCTGATGAGAGTGGTGTGCTGCGAGGCCCATAGGAAGAGCTGG	540
Qy	577	AAGSCCTCCGATGCGTCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGG	636
Db	541	AAGSCCTCCGATGCGTCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGG	600
Qy	637	GCAAAACCGAGGGCCCGCAACTCTTCGAAAACTTGTGTGTGTGCTGAGTCTCTGTGTGGGG	696
Db	601	GCAAAACCGAGGGCCCGCAACTCTTCGAAAACTTGTGTGTGTGCTGAGTCTCTGTGTGGGG	660
Qy	697	CACCCCAACGACACCCCTCTTCCCTCAGGAGGCGCGTGGGCAGACAGGGAGCGCGGGC	756
Db	661	CACCCCAACGACACCCCTCTTCCCTCAGGAGGCGCGTGGGCAGACAGGGAGCGCGGGC	720

Qy	757	TTTGGCCCTGCTGCTGCTCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCC	816
Db	721	TTTGGCCCTGCTGCTGCTCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCC	780
Qy	817	CCCTGCTGCGCCCTGAGAGCGGCTCTGCTGTATCTCAAGAGAGCCCTGTGCCAGGCCG	876
Db	781	CCCTGCTGCGCCCTGAGAGCGGCTCTGCTGTATCTCAAGAGAGCCCTGTGCCAGGCCG	840
Qy	877	TCCACCCCTGGAGTGGTCTTCTTCAGAGCTGTTCCCGAGCCACAGGCGCTGTACGACCC	936
Db	841	TCCACCCCTGGAGTGGTCTTCTTCAGAGCTGTTCCCGAGCCACAGGCGCTGTACGACCC	900
Qy	937	ACGATGTGCCGCAAGCACTGTCTCACCATCCCGCACCCACAGACACACAGCGAGGCTGG	996
Db	901	ACGATGTGCCGCAAGCACTGTCTCACCATCCCGCACCCACAGACACACAGCGAGGCTGG	960
Qy	997	AGTCCAGGCGCACTTTTCAGAGTGTCTTCTTCGCGTGCATGCTGTCTCTCTCTCTCTCT	1056
Db	961	AGTCCAGGCGCACTTTTCAGAGTGTCTTCTTCGCGTGCATGCTGTCTCTCTCTCTCTCT	1020
Qy	1057	TCTCTTCCCGCACTTCTTCTTCTGACCCCTCCCGTCCGCTGGTTCGTATCAAGACT	1116
Db	1021	TCTCTTCCCGCACTTCTTCTTCTGACCCCTCCCGTCCGCTGGTTCGTATCAAGACT	1080
Qy	1117	CCTCAAAACCCCGTCCCGGTGTCTCTGCTGTGTGACAGCTCGCTCTTCTCTCTCTCTCT	1176
Db	1081	CCTCAAAACCCCGTCCCGGTGTCTCTGCTGTGTGACAGCTCGCTCTTCTCTCTCTCTCT	1140
Qy	1177	AGTATCCAAAGGGATGGACCCAGGCTCTGCGGAGGTTCCACCTTGGATCCAGGAAGA	1236
Db	1141	AGTATCCAAAGGGATGGACCCAGGCTCTGCGGAGGTTCCACCTTGGATCCAGGAAGA	1200
Qy	1237	ACCTTCCACCCCTGCTGCTGGGTGGGCAAGGCTACAGGCTGCTTCTCTCTCTCTCTCT	1296
Db	1201	ACCTTCCACCCCTGCTGCTGGGTGGGCAAGGCTACAGGCTGCTTCTCTCTCTCTCTCT	1260
Qy	1297	ACCCCACTGTCTCTCATGTGCCATGGGCTGCTTCCCAAGTGAACCTGCGAAAGTGGAGC	1356
Db	1261	ACCCCACTGTCTCTCATGTGCCATGGGCTGCTTCCCAAGTGAACCTGCGAAAGTGGAGC	1320
Qy	1357	ATCGAGGTAGAGGGGAAACAGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCGCTTCTCT	1416
Db	1321	ATCGAGGTAGAGGGGAAACAGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCGCTTCTCT	1380
Qy	1417	ACCATTTCCCGCACAGAGCTTTTCCCTTGGCTGCGCGCTGCTCTTTGGGGAGAC	1476
Db	1381	ACCATTTCCCGCACAGAGCTTTTCCCTTGGCTGCGCGCTGCTCTTTGGGGAGAC	1440
Qy	1477	TGAGCTCAGAGGAGGAGTGTCTTTCAGAGAAGGAAACAAAATGAGGGGTGGCAGGATAAAA	1536
Db	1441	TGAGCTCAGAGGAGGAGTGTCTTTCAGAGAAGGAAACAAAATGAGGGGTGGCAGGATAAAA	1500
Qy	1537	GTCACTCTCAATCTCTACCTCCCATGAGCATGAACAATAATTTCTCTCACTGGCTCCC	1596
Db	1501	GTCACTCTCAATCTCTACCTCCCATGAGCATGAACAATAATTTCTCTCACTGGCTCCC	1560
Qy	1597	AAATTTAAAGATGTGACCAAGGCGCTGTGGTACTCCAGGGGCAAGGAGCCCTGGGGT	1656
Db	1561	AAATTTAAAGATGTGACCAAGGCGCTGTGGTACTCCAGGGGCAAGGAGCCCTGGGGT	1620
Qy	1657	CAGTGACACTGTTCAGGCGCAACCATGCACTCCAAGAGGGAGCATTTGGAAATGAAGGAC	1716
Db	1621	CAGTGACACTGTTCAGGCGCAACCATGCACTCCAAGAGGGAGCATTTGGAAATGAAGGAC	1680
Qy	1717	TAGCTCTATGTATCAGTTTAAGAGCAAGGAGAGTGGCCAGGACAGAGTTTGACAA	1776
Db	1681	TAGCTCTATGTATCAGTTTAAGAGCAAGGAGAGTGGCCAGGACAGAGTTTGACAA	1740
Qy	1777	GCAGAGGGGAATCTAGCAACAGAGGCGCTCTTAGGCCCATCTTCCATTTCTTAGGTAA	1836
Db	1741	GCAGAGGGGAATCTAGCAACAGAGGCGCTCTTAGGCCCATCTTCCATTTCTTAGGTAA	1800

QY 1837 GAAGAGCAATTCCTCAGACTCCCAAGCGGAGGACTGAGCCCTAGCCTTCAGCAACCAAGGT 1896  
Db 1801 GAAGAGCAATTCCTCAGACTCCCAAGCGGAGGACTGAGCCCTAGCCTTCAGCAACCAAGGT 1860  
QY 1897 TCTCCTGGGACCCAAAGTTTATGGAGAGGCAAGCACTTCATGGGAAGAGAGGAA 1956  
Db 1861 TCTCCTGGGACCCAAAGTTTATGGAGAGGCAAGCACTTCATGGGAAGAGAGGAA 1920  
QY 1957 GGCCCTGGGTAGAAACGCTTGCTGCTGTTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCT 2016  
Db 1921 GGCCCTGGGTAGAAACGCTTGCTGCTGTTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCT 1980  
QY 2017 TGCCCTCTACCTCTGTAGAGCTTTGAGGGTTTGCAACCAACACACTGTGGCTACAGGTGAG 2076  
Db 1981 TGCCCTCTACCTCTGTAGAGCTTTGAGGGTTTGCAACCAACACACTGTGGCTACAGGTGAG 2040  
QY 2077 GGAAGAGACTCCTCTCCAGAGTGCTATGTTTCAGGAGGTTTCTTAACCCCATATGGC 2136  
Db 2041 GGAAGAGACTCCTCTCCAGAGTGCTATGTTTCAGGAGGTTTCTTAACCCCATATGGC 2100  
QY 2137 CCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAACCAAGTAATTTACCAAGTTCTACTG 2196  
Db 2101 CCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAACCAAGTAATTTACCAAGTTCTACTG 2160  
QY 2197 GGGTTCTGCCCCACCGTCCCAAGGTGGCGAGGCTAGGAAGAGGGTCAATTTAAAGCCA 2256  
Db 2161 GGGTTCTGCCCCACCGTCCCAAGGTGGCGAGGCTAGGAAGAGGGTCAATTTAAAGCCA 2220  
QY 2257 CACATTAGCTGACCTGCTGCTGAGCCCAACCAAGAACTGGGTGTGAGTATTCTATC 2316  
Db 2221 CACATTAGCTGACCTGCTGCTGAGCCCAACCAAGAACTGGGTGTGAGTATTCTATC 2280  
QY 2317 AACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGATAAGAACCTCACTTCTTACT 2376  
Db 2281 AACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGATAAGAACCTCACTTCTTACT 2340  
QY 2377 CCTCCAAAGAAAGTGGGGAAGAACCAATCAAACTTTCTCTGACTTACCAACACGAG 2436  
Db 2341 CCTCCAAAGAAAGTGGGGAAGAACCAATCAAACTTTCTCTGACTTACCAACACGAG 2400  
QY 2437 AAAACAGCAGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGTGTGGAA 2496  
Db 2401 AAAACAGCAGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGTGTGGAA 2460  
QY 2497 GCAAGGAGAGCAGCAAGTCTTAATCACTGCTTAATCAGAAAAAGGAGACAGTAACTCT 2556  
Db 2461 GCAAGGAGAGCAGCAAGTCTTAATCACTGCTTAATCAGAAAAAGGAGACAGTAACTCT 2520  
QY 2557 AGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTCTGAAAGATGGGACTCTCTTGGGTATCA 2616  
Db 2521 AGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTCTGAAAGATGGGACTCTCTTGGGTATCA 2580  
QY 2617 AGACTATGCCACATCACACTGGGGCTAGGAAGTAGGTATGCCAGCCCTCAAGTCTGT 2676  
Db 2581 AGACTATGCCACATCACACTGGGGCTAGGAAGTAGGTATGCCAGCCCTCAAGTCTGT 2640  
QY 2677 CTTCCAGCCAGGAGCTTGAGAGAGTTATATTGGGAGTGCTCCAACTCTGTGACCAAGTATT 2736  
Db 2641 CTTCCAGCCAGGAGCTTGAGAGAGTTATATTGGGAGTGCTCCAACTCTGTGACCAAGTATT 2700  
QY 2737 TCAGCTTTCCCTGAAGATCAGGCGAGGGTGCCATTCAATTTCTTCTCTAGCCCCCCTC 2796  
Db 2701 TCAGCTTTCCCTGAAGATCAGGCGAGGGTGCCATTCAATTTCTTCTCTAGCCCCCCTC 2760  
QY 2797 AGGAAAGAGGACTATATTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAA 2856  
Db 2761 AGGAAAGAGGACTATATTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAA 2820  
QY 2857 CAGGATTTCTATAGTGTAGTGGCCCTATCCACAGGGCCATGACTGGGAAAGGTATGGG 2916  
Db 2821 CAGGATTTCTATAGTGTAGTGGCCCTATCCACAGGGCCATGACTGGGAAAGGTATGGG 2880  
QY 2917 AGCAGAGGAGAAATGGGATTTTAGGGTGCAGCTACGCTCACCCCTAAACTTTTGTGGCC 2976

Db 2881 AGCAGAGGAGAAATGGGATTTTAGGGTGCAGCTACGCTCACCTAAACTTTTGTGGCC 2940  
QY 2977 TGGGGCATGCTTTGAGGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTACTCGTCAC 3036  
Db 2941 TGGGGCATGCTTTGAGGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTACTCGTCAC 3000  
QY 3037 CACCTCTGCACTGCTGCTGTTGAGACTCCATCCAGCCCCCAGGACGCCACCTGCTCTGA 3096  
Db 3001 CACCTCTGCACTGCTGCTGTTGAGACTCCATCCAGCCCCCAGGACGCCACCTGCTCTGA 3060  
QY 3097 GCCTCCACTATCTCCCTGTGACGGTGAACTTGTGTACTGTGTCTCGGTCCTATATG 3156  
Db 3061 GCCTCCACTATCTCCCTGTGACGGTGAACTTGTGTACTGTGTCTCGGTCCTATATG 3120  
QY 3157 AATTGTGAGCAGGGTTCATCTATTTTAAACA 3187  
Db 3121 AATTGTGAGCAGGGTTCATCTATTTTAAACA 3151

## RESULT 2

BC014511 2700 bp mRNA linear HTC 11-DEC-2003  
LOCUS Homo sapiens RAB15, member RAS oncogene family, mRNA (cDNA clone  
IMAGE:4866926), with apparent retained intron.

## ACCESSION

BC014511

## VERSION

BC014511.1 GI:15680292

## KEYWORDS

HTC.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 2700)

## AUTHORS

Strausberg, R.L., Reingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

REFERENCE 12477932

2 (bases 1 to 2700)

Strausberg, R.

Direct Submission

Submitted (17-SEP-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NTI-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,





481 AGTGACACTCTCAGGCCAACCATGCACTCCAAAGGGGAGCAATTTGGAAATGAAGACT 540  
1718 AGCTCTCTATGTATCAGGTAAAGACCAAGGAGAGCTGGCCAGGAGACAGCAAGTTTGACAG 1777  
541 AGCTCTCTATGTATCAGGTAAAGACCAAGGAGAGCTGGCCAGGAGACAGCAAGTTTGACAG 600  
1778 CAGAGGGGAATGTAGCAACAGAGGGGCTCTAGGCCCACTCTCCATTTCTTTAGTAAG 1837  
601 CAGAGGGGAATGTAGCAACAGAGGGGCTCTAGGCCCACTCTCCATTTCTTTAGTAAG 660  
1838 AAGAGCAATTTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 1897  
661 AAGAGCAATTTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 720  
1898 CTCCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATCGGGAAGAGAGGAAG 1957  
721 CTCCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATCGGGAAGAGAGGAAG 780  
1958 GCCCTGGGTAGAAACGCTTGGTGTCTCTTTGGGCTTTAGACAAAGGCTCATCTT 2017  
781 GGCCTGNGTAGAAACGCTTGGTGTCTCTTTGGGCTTTAGACAAAGGCTCATCTT 839  
2018 GCCCTCTACCTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 2077  
840 GCCCTCTACCTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 898  
2078 GAAGAGACTCTCTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 2137  
899 GAAGAGACTCTCTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGGTT 955  
2138 CAAGAGTAGCTCGTAGGAGGCGCTTTAAGACGCAAGCAAGTAATTTACCAAGTTCTACT-G 2196  
956 CAAGAAATACCTTTGAAGAGGCGCTTTAAGACGCAAGCAAGTAATTTACCAAGTTCTACTGG 1015  
2197 GGGTTCTCCGCCACCGTCCCAAGGTGGCGGAGGCTTAGGAGAGGGTCAATTTAAAGCC 2255  
1016 GGGTTCTCCGCCACCGTCCCAAGGTGGCGGAGGCTTAGGAGAGGGTCAATTTAAAGCC 1074

RESULT 4  
LOCUS BX397021  
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0DI028YM11 5-PRIME, mRNA sequence.  
BX397021  
BX397021.1 GI:30609314  
EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 956)  
Li.W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7307.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DI028AG06\_D11486\_2&cluster=7307.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/InvitroGen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DI028AG06\_D11486\_2.

FEATURES  
source  
1..956  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI028YM11"

/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and cloned into the NotI and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 27.1%; Score 883.4; DB 13; Length 956;  
Best Local Similarity 98.5%; Pred. No. 46-123; 9; Indels 5; Gaps 5;  
Matches 942; Conservative 0; Mismatches 0

1596 CAAATTTAAAGATGTGGACCAAGGCTGTGGTACTTCCAGGGGCAAGAGAGCCCTGGGG 1655  
2 CCNATTTAAAGATGTGGACCAAGGCTGTGGTACTTCCAGGGGCAAGAGAGCCCTGGGG 60  
1656 TCAGTGACACTGTGACGGCCAAACCATGCACTCCAAAGGGGAGCATTTGGAATGAAGA 1715  
61 TCAGTGACACTGTGACGGCCAAACCATGCACTCCAAAGGGGAGCATTTGGAATGAAGA 120  
1716 CTAGCTCTCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGACAGATTTCAC 1775  
121 CTAGCTCTCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGACAGATTTCAC 180  
1776 AGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTAGGCCCCATCTTCCATTTCTTAGGTA 1835  
181 AGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTAGGCCCCATCTTCCATTTCTTAGGTA 240  
1836 AGAAGAGCATTTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGG 1895  
241 AGAAGAGCATTTCTCAGACTCCAGGCGGAGACTGAGCCTTAGCCTTCAGCAACCAAGG 300  
1896 TTCTCTCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTTCATGGGAGAGAGAGGA 1955  
301 TTCTCTCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTTCATGGGAGAGAGAGGA 360  
1956 AGGCCCTCTGGTGAAGACGCTGTGTCTCTTTGGGCTTTAAGACAAAGCGCTCATC 2015  
361 AGGCCCTCTGGTGAAGACGCTGTGTCTCTTTGGGCTTTAAGACAAAGCGCTCATC 420  
2016 TTGCCCTCTACCTCTGATAGGCTTCAGGTTTGGCCAAACACACTGTGGCTACAGGTGGA 2075  
421 TTGCCCTCTACCTCTGATAGGCTTCAGGTTTGGCCAAACACACTGTGGCTACAGGTGGA 480  
2076 GCGAAGAGGACTCTCTCTCCAGAGTCTATGTTTGGGAGGTTTCTTAAACCCATATGG 2135  
481 GCGAAGAGGACTCTCTCTCCAGAGTCTATGTTTGGGAGGTTTCTTAAACCCATATGG 540  
2136 CCCAAGAGTAGCTCGTAGGAGGCGCTTTAAGAGCGGAACAAGTAATTTACCAAGTTCTACT 2195  
541 CCCAAGAGTAGCTCGTAGGAGGCGCTTTAAGAGCGGAACAAGTAATTTACCAAGTTCTACT 600  
2196 GGGGTTCTCCGCCACCGTCCCAAGGTGGGCGAGGCTTAGGAAGAGGCTCATTTTAAGCC 2255  
601 GGGGTTCTCCGCCACCGTCCCAAGGTGGGCGAGGCTTAGGAAGAGGCTCATTTTAAGCC 660  
2256 ACATTTAGCTGCACTGCGTGGCTGCAAGCCAAACCAAGAACTGGTGTGAGTATTCAT 2315  
661 ACATTTAGCTGCACTGCGTGGCTGCAAGCCAAACCAAGAACTGGTGTGAGTATTCAT 720  
2316 CAATTAAGAACCAAAATTCAGGGGCACTCATATGTGAAGGATAGAAGCTCACTTCTTAC 2375  
721 CAATTAAGAACCAAAATTCAGGGGCACTCATATGTGAAGGATAGAAGCTCACTTCTTAC 780  
2376 TCCTCCAAAGAGAGTGGGAAAGAACCAATCAAAACCTTCTCTCTGACTTACCAACCAAG 2435  
781 TCCTCCAAAGAGAGTGGGAAAGAACCAATCAAAACCTTCTCTCTGACTTACCAACCAAG 840  
2436 GAAACACAGCAGGAGGCTGCTCAGGACTTGAAGGAGGATAGCTTATAGTGGTGA 2495  
841 G-AAACACAGCAGGAGGCTGCTCAGGACTTGAAGGAGGATAGCTTATAGTGGTGA 899  
2496 AGCAAGAGGAGCAGGAGGATGTTGTAATCACTGGCTTATGAGAAAGGAGCAGCTA 2551



```
Db 900 AGC-AAGGAGACGAGGAAGTGT-AATCACTGGCT-ATGAGAAAAAGACAGCTA 952
|||||
BQ937298 957 bp mRNA linear EST 21-AUG-2002
AGENCOURT 10018726 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482195
5', mRNA sequence.
BQ937298
BQ937298.1 GI:22352681
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2664 row: m column: 12
High quality sequence stop: 555.
FEATURES
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6482195"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 26.3%; Score 856; DB 13; Length 957;
Best Local Similarity 97.4%; Pred.No.5.2e-119;
Matches 890; Conservative 0; Mismatches 22; Indels 2; Gaps 2;
Qy 1314 TGTGTCATGGCGCTGCTCCCGAGTACCTGCGAAGTGGAGCATCGAGGTAGAGGGAA 1373
Db 1 TGTGTCATGGCGCTGCTCCCGAGTACCTGCGAAGTGGAGCATCGAGGTAGAGGGAA 60
Qy 1374 ACAGCAACCGGGAGTCTTCAGAGCTGGGGTGCCTACCTACCCATCCCGACAG 1433
Db 61 ACAGCAACCGGGAGTCTTCAGAGCTGGGGTGCCTACCTACCCATCCCGACAG 120
Qy 1434 AGCTTTGGCCCTGCTTGGCTGCGCGCTCTTTGGGGAACGTGAGCTCAGAGCGAGT 1493
Db 121 AGCTTTGGCCCTGCTTGGCTGCGCGCTCTTTGGGGAACGTGAGCTCAGAGCGAGT 180
Qy 1494 GCTTCAGAGAGGAACAAATAGGGGTGGCAGGATAAAAGTACCTCATCTCTCTA 1553
Db 181 GCTTCAGAGAGGAACAAATAGGGGTGGCAGGATAAAAGTACCTCATCTCTCTA 240
Qy 1554 CTCTCCATGCAGCATGAACACAAATTTCTCTCACCTGGCTCCCAATTTAAAGATGTGGA 1613
Db 241 CTCTCCATGCAGCATGAACACAAATTTCTCTCACCTGGCTCCCAATTTAAAGATGTGGA 300
```

```
Qy 1614 CCAAGGCTGTGGTACTTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGC 1673
Db 301 CCAAGGCTGTGGTACTTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGC 360
Qy 1674 CAACCATGCATCCCAAAAGGGAGCAATTTGGAAATGAAGACTAGCTCTCTATGTATCAG 1733
Db 361 CAACCATGCATCCCAAAAGGGAGCAATTTGGAAATGAAGACTAGCTCTCTATGTATCAG 420
Qy 1734 GTTAAGAGCAAGGAGAGCTGGCCAGGAGCAAGCAAGTTTGCACAGAGGGGAATGTAGC 1793
Db 421 GTTAAGAGCAAGGAGAGCTGGCCAGGAGCAAGCAAGTTTGCACAGAGGGGAATGTAGC 480
Qy 1794 AACACGAGGGGCTCTAGGGCCCATCTTCCATTTCTTAGGTAAAGAGCAATTTCTCTCAG 1853
Db 481 AACACGAGGGGCTCTAGGGCCCATCTTCCATTTCTTAGGTAAAGAGCAATTTCTCTCAG 540
Qy 1854 ACTCCAGGGGAGGACTAGAGCTTCCAGCAACCAAGTTCTCTCTGGGACCCAAAG 1913
Db 541 ACTCCAGGGGAGGACTAGAGCTTCCAGCAACCAAGTTCTCTCTGGGACCCAAAG 600
Qy 1914 TTTATGGAGAGGGCAAGACTTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1973
Db 601 TTTATGGAGAGGGCAAGACTTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Qy 1974 CTTGGTGTCTTCTTCTTTGGCCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTCTGA 2033
Db 661 CTTGGTGTCTTCTTCTTTGGCCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTCTGA 720
Qy 2034 TAGGCTTGAGGTTTGGCAACCACTGTGGCTACAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2093
Db 721 TAGGCTTGAGGTTTGGCAACCACTGTGGCTACAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 2094 TCCAGAGTGCTATGTTTCAGGAAGTTCTTTAAACCCATATCGCCACAGAGTAGTCTGAG 2153
Db 781 CTCAGAGTGCTATGTTTCAGGAAGTTCTTTAAACCCATATCGCCACAGAGTAGTCTGAG 839
Qy 2154 GAGGCGCTTTAAAGACGGAACAAGTAATTTACCAAGTTCTACTGGGTTCTCTGC-CCACCG 2212
Db 840 AGAGCCCTTTAAGACGGAACAAGTAATTTACCAAGTTCTACTGGGTTCTCTGC-CCACCG 899
Qy 2213 TCCCAAGGTGGCG 2226
Db 900 TCCCAAGGTGGCG 913
RESULT 6
BQ937020 915 bp mRNA linear EST 13-MAY-2003
LOCUS clone CS0D1028YML1 5-PRIME, mRNA sequence.
DEFINITION BQ937020.1 GI:30609312
ACCESSION BQ937020
VERSION BQ937020.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 915)
Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7307.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1028AG06.D11486.lcluster=7307.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
Faraday Avenue Genoscope sequence ID : CS0D1028AG06_D11486_1.
FEATURES
  source
    Location/Qualifiers
      1..915
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0D1028YM11"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
  Query Match      26.0%; Score 847.2; DB 13; Length 915;
  Best Local Similarity 98.3%; Pred. No. 1.1e-117;
  Matches 882; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
QY 1615 CAAGGCGCTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCACTGACACTGTGACGCC 1674
DB 22 CCAGGCGCTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCACTGACACTGTGACGCC 81
QY 1675 AACCATGCATCTCCAAAGGGAGCACTTTGGAATGAAGGACTAGCTTCCTATGTATCAGG 1734
DB 82 AACCATGCATCTCCAAAGGGAGCACTTTGGAATGAAGGACTAGCTTCCTATGTATCAGG 141
QY 1735 TTAAGACAGGGAGAGCTGCCAGGACAGCAGCTTTGCACAGCAGAGGGGATGTAGCA 1794
DB 142 TTAAGACAGGGAGAGCTGCCAGGACAGCAGCTTTGCACAGCAGAGGGGATGTAGCA 201
QY 1795 ACAGCAGGCGCTCTAGGCGCCCATCTTCCATTTCTTAGSTAAGAGAGCATTTCTCTCAGA 1854
DB 202 ACAGCAGGCGCTCTAGGCGCCCATCTTCCATTTCTTAGSTAAGAGAGCATTTCTCTCAGA 261
QY 1855 CTCCAGGCGGAGACTGAGCCTAGCCTTCAGCAACCAAGTTCTCTGGGACCCAAAGT 1914
DB 262 CTCCAGGCGGAGACTGAGCCTAGCCTTCAGCAACCAAGTTCTCTGGGACCCAAAGT 321
QY 1915 TTATGGAGAGGSCAAAGACTTCATGGGAGAGAGAGGAGGCCCTGGGTAGAAACGC 1974
DB 322 TTATGGAGAGGSCAAAGACTTCATGGGAGAGAGAGGAGGCCCTGGGTAGAAACGC 381
QY 1975 TTGGTGTCTTCTCTTTGGGCTTTAAGCAAAAGCGCTCATCTTGCCCTCTACCTCTCTGAT 2034
DB 382 TTGGTGTCTTCTCTTTGGGCTTTAAGCAAAAGCGCTCATCTTGCCCTCTACCTCTCTGAT 441
QY 2035 AGGCTTGAGGGTTTGCCAAACCACTGTGGCTACAGGTGGAGGAAAGAGACTCCTTCCT 2094
DB 442 AGGCTTGAGGGTTTGCCAAACCACTGTGGCTACAGGTGGAGGAAAGAGACTCCTTCCT 501
QY 2095 CCAGAGTCTATGTTCAGGAGTTTCTTTACCCCATATGCCCCAAGAGTAGCTCGTAGG 2154
DB 502 CCAGAGTCTATGTTCAGGAGTTTCTTTAACCCCATATGCCCCAAGAGTAGCTCGTAGG 561
QY 2155 AGGCGCTTTAAGACGGAACAAGTAATTTACAGTTCTACTGGGGTTCTCTGCCACCGCTC 2214
DB 562 AGGCGCTTTAAGACGGAACAAGTAATTTACAGTTCTACTGGGGTTCTCTGCCACCGCTC 621
QY 2215 CCAAGTGGGCGAGCCCTAGGAAAGGGGTCAATCTTAAGCCACACATTAAGTGTGACTGGG 2274
DB 622 CCAAGTGGGCGAGCCCTAGGAAAGGGGTCAATCTTAAGCCACACATTAAGTGTGACTGGG 681
QY 2275 TGGCTGAGCCCAACAAAGAACTGGGTGTGAGTATTCATCAACTAAGACCAAAATCC 2334
DB 682 TGGCTGAGCCCAACAAAGAACTGGGTGTGAGTATTCATCAACTAAGACCAAAATCC 741
QY 2335 AGGCACTCATATGTGAAGGATAAGAACTCACTTCTTACTCTCTCCAAAGAAAGTGGG 2394
DB 742 AGGCACTCATATGTGAAGGATAAGAACTCACTTCTTACTTCTTCTTCTTCTTCTTCTT 800
QY 2395 GAAAGAACCATCAACCTTCTCTCTGACTTATCCAAACAGGAGAAAAGAGGAGAGGGT 2454
```

```
801 GRAAGAACCATCAACCTTTCTNCTGACTTACCAACACAGG-AAACAGCNAGAGGGT 859
2455 GGCTCAGGACTTAGGACAGGGTATAGCTTAGATGCTGGAAGCAAGAGGAGCAGG 2511
860 GGCTCAGGACTTAGGACAGGGTATAGCTNANATGCT-GAAACAAAGGAGAGCAGG 915

RESULT 7
LOCUS      BQ716774
DEFINITION AGENCOURT 8217893 Lupeki_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6187941 5', mRNA sequence.
ACCESSION  BQ716774
VERSION    BQ716774.1 GI:21855671
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 908)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-remail.nih.gov
          Tissue Procurement: Dr. James R. Lupski
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM13582 row: 1 column: 22
          High quality sequence stop: 638.
FEATURES   Location/Qualifiers
            source
              1..908
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6187941"
                /sex="male"
                /tissue_type="sympathetic trunk"
                /dev_stage="adult, 16 yr"
                /lab_host="DH10B"
                /clone_lib="Lupeki_sympathetic trunk"
                /notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                Directionally cloned using the following adaptors:
                5'-GACTAGTTCCTAGATCGGAGCGGCCGCCCT(15)-3'. Size selected >
                1 kb for average insert length 1.9 kb. This is a primary
                library, non-amplified. Library constructed by Life
                Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor
                College of Medicine); available through Life
                Technologies."
```

```
Query Match      25.8%; Score 840.2; DB 13; Length 908;
  Best Local Similarity 97.6%; Pred. No. 1.3e-116;
  Matches 884; Conservative 0; Mismatches 19; Indels 3; Gaps 3;
QY 1014 GCTGCTCCTTTCTCCGTCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
DB 1 GCTGCTCCTTTCTCCGTCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 1074 CTTTCTCTGACCCCTCCCTCCGTCGTTTCGATCAAGACTCCTCAAAACCCGTCCTCC 1133
DB 61 CTTTCTCTGACCCCTCCCTCCGTCGTTTCGATCAAGACTCCTCAAAACCCGTCCTCC 120
QY 1134 CGTGTGCTCTGTGTGTCAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1193
DB 121 CGTGTGCTCTGTGTGTCAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
```

1194 GACCCAGGCTCGTGGGAGGTTCCACCCCTTGATCCAGAGAAACCCCTCCACCCCTGCTC 1253  
181 GACCCAGGCTCGTGGGAGGTTCCACCCCTTGATCCAGAGAAACCCCTCCACCCCTGCTC 240  
1254 GTGGGTGGCCAAAGGCTACAGGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1313  
241 GTGGGTGGCCAAAGGCTACAGGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
1314 TGTGCCATGGCCCTGCTCTCCCACTGACCTCCGAAGTGGAGCATCAGAGTAGAGGGAA 1373  
301 TGTGCCATGGCCCTGCTCTCCCACTGACCTCCGAAGTGGAGCATCAGAGTAGAGGGAA 360  
1374 ACAGCAACCGGGAGTCTCTGAGGCTGGGGCTGCGCTTACCTCTACCAATCCCGGACAG 1433  
361 ACAGCAACCGGGAGTCTCTGAGGCTGGGGCTGCGCTTACCTCTACCAATCCCGGACAG 420  
1434 AGCTTTCCTCTGCTGGCTGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1493  
421 AGCTTTCCTCTGCTGGCTGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
1494 GCTTCAGAGAGGAAACAAATAGGGGTGGCAGGATAAAGTCACTCTCATTTCTCTA 1553  
481 GCTTCAGAGAGGAAACAAATAGGGGTGGCAGGATAAAGTCACTCTCATTTCTCTA 540  
1554 CTTCCCATGACATGAACAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGA 1613  
541 CTTCCCATGACATGAACAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGA 600  
1614 CCAAGGCTGTGGTACTCTCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTGTGAGC 1673  
601 CCAAGGCTGTGGTACTCTCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTGTGAGC 660  
1674 CAACATGCACTCCACAAGGGGAGCATTTGGAATGAAGACTAGCTTCTATGTATCAG 1733  
661 CAACATGCACTCCACAAGGGGAGCATTTGGAATGAAGACTAGCTTCTATGTATCAG 720  
1734 GTTAAAGCAAGGAGAGCTGGCCAGGACAGCATTTTGACAG-CAGAGGGGAATGTAG 1792  
721 GTTAAAGCAAGGAGAGCTGGCCAGGACAGCATTTTGACAGCAGAGGGGGATGTAG 780  
1793 CAACAGCA-GGGCTCTAGGCCCATCTTCATTTCTAGGTAAGAGACATTTCTC 1851  
781 CAACAGCAGGGGCTCTAGGCCCATCTTCATTTCTAGTAAAGAGACATTTCTC 840  
1852 AGACTCCAGCGGAGGACTCAGCTAGCC-TTACAGCAACCAAGTCTCTCTGGGACCA 1910  
841 AGACTCCAGCGGAGGAAATGAACCTAGCCCTTACAGCAACCAAGTCTCTCTGGGAC 900  
1911 AAGTTT 1916  
901 AAGTT 906

RESULT 8  
BX326279  
LOCUS  
DEFINITION BX326279 Homo sapiens PLACENTA COT 25-NORMALIZED EST 02-MAY-2003  
clone CS0D1028YM11 5-PRIME, mRNA sequence.  
ACCESSION  
VERSION BX326279.1 GI:30336558  
KEYWORDS EST.  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 918)  
TITLE Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
JOURNAL Full-length cDNA libraries and normalization  
COMMENT Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7307.r For more information about this cluster, see [http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAA0152B10\\_CS01419\\_1&cluster=7307.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAA0152B10_CS01419_1&cluster=7307.r). Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : [http://fulllength.invitrogen.com/InvitrogenCorporation1600FaradayAvenueGenoscopeSequenceID:CS0BAA0152B10\\_CS01419\\_1](http://fulllength.invitrogen.com/InvitrogenCorporation1600FaradayAvenueGenoscopeSequenceID:CS0BAA0152B10_CS01419_1). Location/Qualifiers  
1. 918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1028YM11"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES  
source

ORIGIN  
Query Match 25.7%; Score 837.8; DB 13; Length 918;  
Best Local Similarity 97.9%; Pred. No. 2.9e-116;  
Matches 848; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 380 AGGCGTCCAGAGATCCTTATTGGGAATAAGCTGTAGGAGCAGAAACGGCAGGTGG 439  
Db  
Qy 52 AGGCGTCCAGAGATCCTTATTGGGAATAAGCTGTAGGAGCAGAAACGGCAGGTGG 111  
Db  
Qy 440 AAGAGACAAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAAAACAAGTGC 499  
Db 112 AAGAGACAAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAAAACAAGTGC 171  
Qy 500 CTGACCAACCTCAACATTAAGAGTCAATCACCGCTCTGACAGAGCTGGTCTGCAGGC 559  
Db 172 CTGACCAACCTCAACATTAAGAGTCAATCACCGCTCTGACAGAGCTGGTCTGCAGGC 231  
Qy 560 CCATAGGAAGGAGCTGGAAGGCTCCGGATGCGTGCCAGCAATGAGTTGGCACTGGCAGA 619  
Db 232 CCATAGGAAGGAGCTGGAAGGCTCCGGATGCGTGCCAGCAATGAGTTGGCACTGGCAGA 291  
Qy 620 GCTGGAGGAGGAGGAGGCGCAAAACCGAGGGCCAGCGAACTCTTCGAAAACCTGCTGTG 679  
Db 292 GCTGGAGGAGGAGGAGGCGCAAAACCGAGGGCCAGCGAACTCTTCGAAAACCTGCTGTG 351  
Qy 680 CTGAGTCTGTGGGGACCCCAACAGACACCCCTCTTCCCTCAGAGGCGCGTGGGCA 739  
Db 352 CTGAGTCTGTGGGGACCCCAACAGACACCCCTCTTCCCTCAGAGGCGCGTGGGCA 411  
Qy 740 GACAGGGGAGCGGGGCTTTGCCCTGTGCTCTCTCGTGTGATGACCTATTGAGTA 799  
Db 412 GACAGGGGAGCGGGGCTTTGCCCTGTGCTCTCTCGTGTGATGACCTATTGAGTA 471  
Qy 800 TCAGTAGCCACTACTCCCTCCCTGCGCTTGCGCTTGAGAGGGCTCTGCTGTCTCATCTCAAGCAG 859  
Db 472 TCAGTAGCCACTACTCCCTCCCTGCGCTTGCGCTTGAGAGGGCTCTGCTGTCTCATCTCAAGCAG 531  
Qy 860 CCCCTGTCCCAGCCCGCTCAACCTCGAGTGGTCTTTTTCAGCTGCTTTTCCCGACCCACA 919  
Db 532 CCCCTGTCCCAGCCCGCTCAACCTCGAGTGGTCTTTTTCAGCTGCTTTTCCCGACCCACA 591  
Qy 920 GGCTGTGTAGACCCCGACGATGCGCGAGCACTGTCTCACCATCCCGACCCACACAG 979  
Db 592 GGCTGTGTAGACCCCGACGATGCGCGAGCACTGTCTCACCATCCCGACCCACACAG 651  
Qy 980 ACAACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTCTCCGTCGATCGTGT 1039  
Db 652 ACAACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTCTCCGTCGATCGTGT 711  
Qy 1040 CTCCTCTCTGCTTTTCTCTCTTCCCGACCTTCTCTTCTGACCCCTCCCTCCCGTG 1099  
Db 712 CTCCTCTCTGCTTTTCTCTCTTACCCCACTTCTCTTCTCTGACCCCTCCCTCCCGTG 771

```

QY 1100 CGTTTCGTCAAGCTCCTCAAAACCCCGTCCCGTGTGTCTGTCTGTGTGAGCTCGC 1159
Db 772 CGTGACGTATCAAGAGTCTCTCAAAACCCCGTACCCCGTGTGTACTGTGTGTGAGCTCGC 831
QY 1160 TCATTTCCTTCTTCCCTAAGCTATCCAAAGGGGATGGACCCAGGCTCGTGGGGAGGTTCCAC 1219
Db 832 TCATTTCCTTCTTCCCTAAGCTATCCAAAGGGGATGGACCCAGGCTCGTGGGGAGGTTTCAC 891
QY 1220 CTTTGGATCCAGGAAGAACCTCTCCAC 1245
Db 892 CTTTGGATCCAGGAAGAACCTGCCCC 917

RESULT 9
BQ220195 901 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7594485 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6061422
5', mRNA sequence.
ACCESSION BQ220195.1 GI:20401595
VERSION BQ220195.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 901)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1331 row: e column: 07
High quality sequence stop: 606.
Location/Qualifiers
1. 901
/organism="Homo sapiens"
/mo.-type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6061422"
/tissue type="melanotic melanoma"
/lab host="DH10B (phage-resistant)"
/clone.lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 Kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 24.9%; Score 810.6; DB 13; Length 901;
Best Local Similarity 98.7%; Pred. No. 3.5e-112;
Matches 838; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 2160 CTTTAAAGACGGAAACAAGTAATTTACCAAGTTCTACTGGGGTTCCTGCCACCGTCCCAAG 2219
Db 1 CTTTAAAGACGGAAACAAGTAATTTACCAAGTTCTACTGGGGTTCCTGCCACCGTCCCAAG 60

QY 2220 GTGGCGAGGCTTAGAGAGAGGCTCATTTCTTAAGCCACACATTAGCTGCATGGGTGCT 2279
Db 61 GTGGCGAGGCTTAGAGAGAGGCTCATTTCTTAAGCCACACATTAGCTGCATGGGTGCT 120

QY 2280 GCAGCCAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAATCCAGGGC 2339
Db 121 GCAGCCAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAATCCAGGGC 180

QY 2340 ACTCATATGTGAGGATAGAAACCTCACTTCTTACTCTCTCCAAAAAGAGTGGGGAAAG 2399

```

```

Db 181 ACTCATATGTGAGGATAGAAACCTCACTTCTTACTCTCCAAAAAGAGTGGGAAAG 240
QY 2400 AACCATCAAAACCTTTCTCTCTCACTTACCAAAACAGGAAACAGCAGGAGAGGTTGGCTC 2459
Db 241 AACCATCAAAACCTTTCTCTCTCACTTACCAAAACAGGAAACAGCAGGAGAGGTTGGCTC 300
QY 2460 AGGACTTAGGGACAGGGTATAGCTTAGATGTGTGAAAGCAAAAGGAGAGCAGGAAGTTGTA 2519
Db 301 AGGACTTAGGGACAGGGTATAGCTTAGATGTGTGAAAGCAAAAGGAGAGCAGGAAGTTGTA 360
QY 2520 AATCACTGGCTAATGAGAAAAGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTG 2579
Db 361 AATCACTGGCTAATGAGAAAAGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTG 420
QY 2580 GAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAACACCTATGCCACATCACACTGG 2639
Db 421 GAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAACACCTATGCCACATCACACTGG 480
QY 2640 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGT 2699
Db 481 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGT 540
QY 2700 TATATTGGGAGTGGCTCCAATCTGTGGACCACTATTTTCACTTTCCCTGAAGATCAGGC 2759
Db 541 TATATTGGGAGTGGCTCCAATCTGTGGACCACTATTTTCACTTTCCCTGAAGATCAGGC 600
QY 2760 AGGTCGCCATTCATTCCTCTCTCTAGCCCTCAGAAAGAGGAGTATATTTGTA 2819
Db 601 AGGTCGCCATTCATTCCTCTCTAGCCCTCAGAAAGAGGAGTATATTTGTA 660
QY 2820 CTGTACCTTAGGGTTCCTGGAAGGAAACATGAATCAGGATTTCTATAGCTGATAGGC 2879
Db 661 CTGTACCTTAGGGTTCCTGGAAGGAAACATGAATCAGGATTTCTATAGCTGATAGGC 720
QY 2880 CCTATCCAAAGGGCCATGATCTGGGAAAGGATAGGAGCAG-AAGGAGNAATTTGGATTT 2938
Db 721 CCTATCCAAAGGGCCATGATCTGGGAAAGGATAGGAGCAGAAAGGAGAAATTTGGGATTT 780
QY 2939 TAGGGTCAGCTACGCTACCCCTAACTTTTGTGGCT-GGGGCATGTCTTGGAGGCCCA 2997
Db 781 TAGGGTCAGCTACGCTACCCCTAACTTTTGGGGCTGGGGGCATGTCTTGGAGGCCCA 840
QY 2998 GACTGTAA 3006
Db 841 GACTGTAA 849

RESULT 10
BQ348287 929 bp mRNA linear EST 05-MAY-2003
LOCUS BQ348287 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI028YM11 5-PRIME, mRNA sequence.
ACCESSION BQ348287
VERSION BQ348287.1 GI:30369275
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 929)
L1,W.B., Gruber,C., Jesse,J., Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7307.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF024ZB10_AF02283_2&cluster=7307.r.

```

907	QY	TTCCCCAGCCACAGGCTGCTACGACCCCCACGATGTCGCGAAGCACTGTCTCACCATC	966
543	Db	TTCCCCAGCCACAGGCTGCTACGACCCCCACGATGTCGCGAAGCACTGTCTCACCATC	602
967	QY	CCGACACCCACAGACAACAGCCAGGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTTCT	102
603	Db	CCGACACCCACAGACAACAGCCAGGGCTGGAGTCCAGGCCACTTTTCAGCTGCTCTTTCT	662
1027	QY	CCGTGCACTCGTGTCTTCTCTCTGCTTTTCTCTCTTCCGCCACTTCTCTTTCTCTGACCC	108
663	Db	CCGTGCACTCGTGTCTTCTCTCTGCTTTTCTCTCTT - CCCCACCTCTCTTTCTCTGACCC	721
1087	QY	CTCCCCCTCCGGTGGCTTTTCGATCAAAAGCTCTCTAAACCCCGTGCCCGGTGTCCTGCT	114
722	Db	CT - CCCCCTCCGGTGGCTTTTCGAT - AAAGCTCTCTAAACCCCGGTGTCCTGCTGCT	778

	Query Match	24.6%	Score 801.6	DB 12	Length 924
	Best Local Similarity	93.7%	Pred. No. 7.8e-111		
	Matches 869	Conservative 0	Mismatches 54	Indels 4	Gaps 3
Qy	1504	AGGAACAAATGAGGGGTGGCAGGGATAAAAGTCACTCCATTCTTACCTCCCATGC	1563		
Db	2	AGGAACAAATGAGGGGTGGCAGGGATAAAAGTCACTCCATTCTTACCTCCCATGC	61		
Qy	1564	AGCATGAACAATTTCTTCCACCTGGCTCCCAAATTTAAAGATGTGGACCAAGGCCTG	1623		
Db	62	AGCATGAACAATTTCTTCCACCTGGCTCCCAAATTTAAAGATGTGGACCAAGGCCTG	121		
Qy	1624	TGGGTACTTCAGGGGCAAGAGAGCCCTGGGGTCAGTGCACCTGTCAAGGCCAACCATGCA	1683		

Db 122 TGGGTACTCCAGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTCAGGCCAACCATGCA 181  
Qy 1684 CTCACAAAGGGGAGGACATTTGGAAATCAAGACTAGCTCTCTATGTATCATGTTTAAGAGCA 1743  
Db 182 CTCACAAAGGGGAGGACATTTGGAAATCAAGACTAGCTCTCTATGTATCATGTTTAAGAGCA 241  
Qy 1744 AGGAGAGCTGGCCAGGAGCAGCAGTGTTCAGCAGCAGAGGGGAATGTAGCAACAGCAGGG 1803  
Db 242 AGGAGAGCTGGCCA-GGACAGCAGTGTTCAGCAGCAGAGGGGAATGTAGCAACAGCAGGG 300  
Qy 1804 CTCCTAGGCCCCATCTTCATTTCTAGTGAAGAGAGATTTCTCTCAGACTCCCAAGC 1863  
Db 301 CTCCTAGGCCCCATCTTCATTTCTAGTGAAGAGAGATTTCTCTCAGACTCCCAAGC 360  
Qy 1864 GGAGGACTGAGCCTAGCCTTCAGCAACCAAGTCTCTCTGGGACCCAAAGTTTATGGGAG 1923  
Db 361 GGAGGACTGAGCCTAGCCTTCAGCAACCAAGTCTCTCTGGGACCCAAAGTTTATGGGAG 420  
Qy 1924 AAGGGCAAGACTTCATGGGAAGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGGTCTG 1983  
Db 421 AAGGGCAAGACTTCATGGGAAGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGGTCTG 480  
Qy 1984 TTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCCCTCTACCTCTGTATAGCTTTGAG 2043  
Db 481 TTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCCCTCTACCTCTGTATAGCTTTGAG 540  
Qy 2044 GGTTCGCAACCACTGTGCTACAGTGGGAGAGAGACTCTCTCTCCAGAGTGC 2103  
Db 541 GGTTCGCAACCACTGTGCTACAGTGGGAGAGAGACTCTCTCTCCAGAGTGC 600  
Qy 2104 TATGTTCAGGAAGTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGAGGCCCTTT 2163  
Db 601 TATGTTCAGGAAGTTCTTTAAACCC-ATATGGCCAGAGTAGCTCGTAGAGGCCCTTT 658  
Qy 2164 AAAGACGGAACAAGTAATTTACCAAGTTCTACTG-GGGTTCTGCCCCACCGTCCCAAGGTG 2222  
Db 659 AAAGACGGAACAAGTAATTTACCAAGTTCTACTGAGGGTCTCTGGGCCACCGTCCCAAGGTG 718  
Qy 2223 GGCAGGCTTAGGAAGGGTCACTCTTAAGCCACATTAAGTCTGCTGGTGGCTGCA 2282  
Db 719 GGCAGGCTTAGGAAGGGTCACTCTTAAGCCACATTAAGTCTGCTGGTGGCTGCA 778  
Qy 2283 GCCAAACAAAGACTGGGTGTTGAGTATTCATCAACTAAGAAACCAAAATCCAGGGCACT 2342  
Db 779 GCCAAACAGAGACTGGGTGTTGAGTATTCATCAAGTAAGAAACCAAAATCCAGGGCACT 838  
Qy 2343 CATATGGAAGGATAAGAACTCTCTTACTCTTCTCCAAAGAAAGAGTGGGGAAGAAC 2402  
Db 839 CATATGGAAGGATAAGAACTCTTCTTATCTCTCCAAAGAAAGTGGGGAAGAAC 898  
Qy 2403 CATCAACCTTTCTCTGACTTACC 2428  
Db 899 ATTAACCTTACTCTGATTACAAAGC 924

RESULT 12  
BQ277739 1046 bp mRNA linear EST 07-MAY-2002  
LOCUS AGENCOURT 7047903 NIH\_MGC\_109 Homo sapiens cdna clone IMAGE:5804597  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ277739  
VERSION BQ277739.1 GI:20487947  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1046)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsabbs-r@mail.nih.gov

Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2042 row: h column: 06  
High quality sequence stop: 570.

FEATURES  
Location/Qualifiers  
1..1046  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5804597"  
/tissue\_type="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/notes="Organ: ovary; Vector: pOTB7; Site: 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G) Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 24.5%; Score 799.2; DB 13; Length 1046;  
Best Local Similarity 96.9%; Pred. No. 1.7e-110;  
Matches 847; Conservative 0; Mismatches 23; Indels 4; Gaps 3;  
Qy 1547 TTCTCTACCTCCCATGCAGCATCAACACAAATTTCTCTCCACCTGGCTCCCAAAATTTAAAG 1606  
Db 1 TTCTCTACCTCCCATGCAGCATCAACACAAATTTCTCTCCACCTGGCTCCCAAAATTTAAAG 60  
Qy 1607 ATGTGGACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGGTCAAGTACACT 1666  
Db 61 ATGTGGACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGGTCAAGTACACT 120  
Qy 1667 GTCAAGCCACCATGCACTCCAAAGGGGAGCATTTGGAAATGAGGACTAGCTCTAT 1726  
Db 121 GTCAAGCCACCATGCACTCCAAAGGGGAGCATTTGGAAATGAGGACTAGCTCTAT 180  
Qy 1727 GTATCAGTTTAAGCAAGGGAGAGCTGGCCAGGCAAGAGAGTTCACAGCAGAGGGGA 1786  
Db 181 GTATCAGTTTAAGCAAGGGAGAGCTGGCCAGGCAAGAGAGTTCACAGCAGAGGGGA 240  
Qy 1787 ATGTAGCAACAGCAGGGCCCTCTTAGGCCCCCATCTTCCATTTCTTTAGTAAGAGAGCAT 1846  
Db 241 ATGTAGCAACAGCAGGGCCCTCTTAGGCCCCCATCTTCCATTTCTTTAGTAAGAGAGCAT 300  
Qy 1847 TCCTCAGACTCCAGGGGAGGACTGAGCCCTAGGCTTCAGCAACCAAGTTCTCTCTGGA 1906  
Db 301 TCCTCAGACTCCAGGGGAGGACTGAGCCCTAGGCTTCAGCAACCAAGTTCTCTCTGGA 360  
Qy 1907 CCCAAAGTTTATGGGAAGGGCAAGAGACTTCATGGGAAGAGAGAGAGCCCTGGGT 1966  
Db 361 CCCAAAGTTTATGGGAAGGGCAAGAGACTTCATGGGAAGAGAGAGAGCCCTGGGT 420  
Qy 1967 AGAAACGCTTTGGTCTGTCTCTTTGGCCCTTTTAAGACAAAGCGCTCATCTTGGCCCTCTAC 2026  
Db 421 AGAAACGCTTTGGTCTGTCTCTTTGGCCCTTTTAAGACAAAGCGCTCATCTTGGCCCTCTAC 480  
Qy 2027 CTCCTGATAGCTTAGGGTTTGGCAACCACTGTGGCTACAGCTGGAGGAGAGGAC 2086  
Db 481 CTCCTGATAGCTTAGGGTTTGGCAACCACTGTGGCTACAGCTGGAGGAGAGGAC 540  
Qy 2087 TCCTTCTCTCAGAGTGTCTATGTTTCCAGGAAGTTTCTTTAAACCCCATATGCGCCCAAGAGTAG 2146  
Db 541 TCCTTCTCTCAGAGTGTCTATGTTTCCAGGAAGTTTCTTTAAACCCCATATGCGCCCAAGAGTAG 600  
Qy 2147 CTCGTAGGAGGCCCTTTTAAAGACGGAACAAAGTAATTTACAGTTCTTACTGGGGTCTCTGC 2206





Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2831 row: j column: 04  
High quality sequence stop: 668.  
Location/Qualifiers  
1. .820  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6598564"  
/tissue\_types="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

FEATURES  
source

Query Match 24.3%; Score 792.2; DB 13; Length 820;  
Best Local Similarity 99.0%; Pred. No. 2.1e-109;  
Matches 797; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

## ORIGIN

QY 2435 GGAACACAGCAGGAGGGTGGCTCAGGACTTAGGGACGGTATAGCTAGTGGGA 2494  
DB 1 GGAACACAGGAGGGTGGCTCAGGACTTAGGGACGGTATAGCTAGTGGGA 60  
QY 2495 AAGCAAGGAGAGCAGGAAGTTGTAATCACTGGCTAATGAGAAAGGACAGCTAACT 2554  
DB 61 AAGCAAGGAGAGCAGGAAGTTGTAATCACTGGCTAATGAGAAAGGACAGCTAACT 120  
QY 2555 CTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTAT 2614  
DB 121 CTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTAT 180  
QY 2615 CAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGGTGTCAGGACCTCAAGTCT 2674  
DB 181 CAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGGTGTCAGGACCTCAAGTCT 240  
QY 2675 GTCTTCAGCCAGGACTTGAAGATTATATTGGGCACTGGCTCCAATCTGTGACCAAGTA 2734  
DB 241 GTCTTCAGCCAGGACTTGAAGATTATATTGGGCACTGGCTCCAATCTGTGACCAAGTA 300  
QY 2735 TTTCAGCTTCCCTGAAGATCAGCAGGGTGCATTCATTGTCTTTCTCTCTAGCCCCC 2794  
DB 301 TTTCAGCTTCCCTGAAGATCAGCAGGGTGCATTCATTGTCTTTCTCTCTAGCCCCC 360  
QY 2795 TCAGGAAGAGGACTATATTGTACTGTACCTAGGGTCTTGGAGGGAACATGGA 2854  
DB 361 TCAGGAAGAGGACTATATTGTACTGTACCTAGGGTCTTGGAGGGAACATGGA 420  
QY 2855 ATCAGGATTCCTATAGACTGATAGGCCCTATCCCAAGGGCCATGACTGGGAAAGGTATG 2914  
DB 421 ATCAGGATTCCTATAGACTGATAGGCCCTATCCCAAGGGCCATGACTGGGAAAGGTATG 480  
QY 2915 GGAGCAGAAGAGAAATTTAGGGTGCAGCTACGCTACCCCTAAACTTTTGTGG 2974  
DB 481 GGAGCAGAAGAGAAATTTAGGGTGCAGCTACGCTACCCCTAAACTTTTGTGG 540  
QY 2975 CCTGGGCGATGCTTGTAGGCGCCAGCTGTTAAGCAGGCTCTGTGGCCTGTTTACTGTC 3034  
DB 541 CCTGGGCGATGCTTGTAGGCGCCAGCTGTTAAGCAGGCTCTGTGGCCTGTTTACTGTC 600  
QY 3035 ACCACCTCTGCACCTGTGCTTTGAGACTCCATCCAGCCCCCAGGACGACCTGTCTCT 3094

DB 601 ACCACCTCTGCACCTGTGCTTTGAGACTCCATCCAGCCCCCAGGACGACCTGTCTCT 660  
QY 3095 GAGCCTCCACTATCTCCCTGTGAGGGTGAACCTCGTGTACTGTCTCGGTCCATATA 3154  
DB 661 GAGCCTCCACTATCTCCCTGTGAGGGTGAACCTCGTGTACTGTCTCGGTCCATATA 720  
QY 3155 TGAATTGTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATT 3214  
DB 721 TGAATTGTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATT 780  
QY 3215 CAAACACCAAAAAA 3239  
DB 781 TCAACACCAAAAAA 805  
RESULT 15  
BU557274  
LOCUS BU557274 863 bp mRNA linear EST 16-SEP-2002  
DEFINITION AGENCOURT 10253405 NIH\_MGC\_109 Homo sapiens cDNA clone  
IMAGE:6585115 5', mRNA sequence.  
ACCESSION BU557274  
VERSION BU557274.1 GI:22907570  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 863)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-femail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2796 row: i column: 19  
High quality sequence stop: 679.  
Location/Qualifiers  
1. .863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6585115"  
/tissue\_types="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

Query Match 24.3%; Score 792.2; DB 13; Length 863;  
Best Local Similarity 97.5%; Pred. No. 2.1e-109;  
Matches 826; Conservative 0; Mismatches 18; Indels 3; Gaps 2;

QY 1547 TTCTCTACCTCCCATGAGCATGAACAAATTTCTCTCCACCTGGCTCCCAAAATTAAAG 1606  
DB 1 TTCTCTACCTCCCATGAGCATGAACAAATTTCTCTCCACCTGGCTCCCAAAATTAAAG 60  
QY 1607 ATGTGACCAAGCGCTGTGGTGTCTCCAGGGCAAGAGAGCCCTGGGGTCAGTGACACT 1666  
DB 61 ATGTGACCAAGCGCTGTGGTGTCTCCAGGGCAAGAGAGCCCTGGGGTCAGTGACACT 120



Qy	1667	GTCTAGGCGCAACCATGCACTCCACAAAGGGGAGCATTTGGAAATGAAGGACTAGCTCTCTAT	1726
Db	121	GTCTAGGCGCAACCATGCACTCCACAAAGGGGAGCATTTGGAAATGAAGGACTAGCTCTCTAT	180
Qy	1727	GTATCAGGTTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGCAGTTTGCACAGCAGAGGGGA	1786
Db	181	GTATCAGGTTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGCAGTTTGCACAGCAGAGGGGA	240
Qy	1787	ATGTAGCAACAGCAGGCGCTCTAGGCGCCCATCTTCCATTTCTTAGTTAAGAGGCAATT	1846
Db	241	ATGTAGCAACAGCAGGCGCTCTAGGCGCCCATCTTCCATTTCTTAGTTAAGAGGCAATT	300
Qy	1847	TCCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTTCAGCAACCAAGGTTCTCTGGGA	1906
Db	301	TCCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTTCAGCAACCAAGGTTCTCTGGGA	360
Qy	1907	CCCAAGTTTATGGAGAGAGGGCAAGACTTTCATGGGAGAGAGAAAGAGGCGCTGGGT	1966
Db	361	CCCAAGTTTATGGAGAGAGGGCAAGACTTTCATGGGAGAGAGAAAGAGGCGCTGGGT	420
Qy	1967	AGAAACGCTTGGTGTCTTCTTTGGCCTTTAAGACAAAGCGCTCATCTTGGCCCTCTAC	2026
Db	421	AGAAACGCTTGGTGTCTTCTTTGGCCTTTAAGACAAAGCGCTCATCTTGGCCCTCTAC	480
Qy	2027	CTCCTGATAGGCTTGGGGTTTGGCAACCACTGTGGCTACAGTGGAGGGAAGAGGAC	2086
Db	481	CTCCTGATAGGCTTGGGGTTTGGCAACCACTGTGGCTACAGTGGAGGGAAGAGGAC	540
Qy	2087	TCCTTCTCTCCAGAGTGCTATGTTTCAGGAAGTTTCTTTAACCCCATATGGCCCAAGAGTAG	2146
Db	541	TCCTTCTCTCCAGAGTGCTATGTTTCAGGAAGTTTCTTTAACCCCATATGGCCCAAGAGTAG	600
Qy	2147	CTCGTAGAGGCCCTTTAAAGACGGAACAAGTAATTTACCAAGTTCTACTGGGGTTCTTGC	2206
Db	601	CTCGTAGAGGCCCTTTAAAGACGGAACAAGTAATTTACCAAGTTCTACTGGGGTTCTTGC	660
Qy	2207	CCACGTCCTCAAGTGGGCGAGGCTAGGAAGGGTCATTTCTTAAGCCACACATTTAGCT	2266
Db	661	CCACGTCCTCAAGTGGGCGAGGCTAGGAAGGGTCATTTCTTAAGCCACACATTTAGCT	720
Qy	2267	GCACGTGGCTGTCAGGCCAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAAC	2326
Db	721	GCACGTGGCTGTCAGGCCAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAAC	780
Qy	2327	CAAAATCCA-GGGCACTCATATGTGAA--GGATAAGAACCTCACTTCTTACTCTCCAA	2383
Db	781	CAAAATCCAAGGGGCACTCATATGTGAAAGGATAAGAAACCTCACTTCTTACTCCCTC	840
Qy	2384	AAAGAAG 2390	
Db	841	CAAAAAG 847	

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:08:42 ; Search time 60 seconds

(without alignments)  
998.335 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGRKEGPGANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	5	ABP62882 Human pol
2	1105	100.0	212	6	AAE29096 Human Ras
3	1105	100.0	401	4	AAU17136 Novel sig
4	1105	100.0	401	7	ADB93844 Human nov
5	1092	98.8	218	3	AAE41604 Human ORF
6	1077	97.5	212	7	ADE58127 Rat Prote
7	832	75.3	188	4	AAU17555 Novel sig
8	832	75.3	188	7	ADB94263 Human nov
9	545.5	49.5	204	4	ABE70670 Drosophil
10	545	49.3	207	7	ADE58121 Human Pro
11	545	49.3	213	4	ABE11916 Human Rab
12	540	48.9	221	5	ABP41333 Human ova
13	532	48.1	203	7	ADB83429 Human Pro
14	532	48.1	203	7	ADB8125 Human Pro
15	532	48.1	246	3	ABE58196 Lung canc
16	530.5	48.0	200	3	ABO9979 Human Rab
17	530.5	48.0	200	3	ABE19165 Amino aci
18	530.5	48.0	200	4	ABE95340 Human pro
19	530.5	48.0	218	3	ABE56993 Human pro
20	529.5	47.9	218	4	ABG07266 Novel hum
21	528.5	47.8	207	4	ABE71647 Drosophil
22	528.5	47.8	207	4	ABG67154 Amino aci
23	528.5	47.8	207	4	ABE92628 Human pro
24	528.5	47.8	207	5	ABP65204 Hypoxia-r
25	526	47.6	201	3	AAE09982 Canine Ra

26	523	47.3	199	3	AAE09980 Human Rab
27	523	47.3	199	3	AAE09981 Human Rab
28	523	47.3	206	6	ABJ26542 Aspergill
29	523	47.3	206	6	ABJ25583 Aspergill
30	519	47.0	216	3	AAE53945 Arabidops
31	519	47.0	216	3	AAE08688 Arabidops
32	519	47.0	253	3	AAE53944 Arabidops
33	519	47.0	254	3	AAE08687 Arabidops
34	514	46.5	201	6	AAO19956 C elegans
35	513.5	46.5	215	3	AAE35215 Zea mays
36	512	46.3	190	6	AAO26373 Ras-like
37	510	46.2	218	3	AAE19220 Arabidops
38	507	45.9	216	3	AAE08006 Arabidops
39	506.5	45.8	209	4	AAE23365 Novel hum
40	501	45.3	224	3	AAE47826 Arabidops
41	501	45.3	234	3	AAE47825 Arabidops
42	501	45.3	335	4	ABG23366 Novel hum
43	479.5	43.4	256	6	AAO15989 Human Ras
44	479.5	43.4	256	7	ADA09336 Human GTP
45	478.5	43.3	201	2	AAE00919 Human Rab

#### ALIGNMENTS

##### RESULT 1

ABP62882  
ID ABP62882 standard; protein; 212 AA.

AC ABP62882;

DT 14-OCT-2002 (first entry)

DE Human polypeptide SEQ ID NO 319.

XX Human; vulnery; dermatological; neuroprotective; nontropic; cancer;  
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;  
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
KW burn; central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; immune disorder;  
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.

OS Homo sapiens.

PN WO200218424-A2.

PD 07-MAR-2002.

XX 31-AUG-2001; 2001WO-US027093.

XX 01-SEP-2000; 2000US-00654935.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;

PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;

XX WPI; 2002-583321/62.

DR N-PSDB; ABQ93361.

XX New polynucleotide and polypeptides, useful for treatment and diagnosis  
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral  
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple  
PT sclerosis, diabetes and allergies.

XX Claim 20; SEQ ID NO 319; 284pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising one of  
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising  
CC administering to a mammalian subject a composition comprising the protein  
CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).  
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.  
CC (I) is useful for gene therapy of diseases and (II) can be used for

CC therapeutic treatment. Diseases that may be treated include wound healing  
 CC and tissue repair, burns, central nervous system disorders (e.g.  
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral  
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple  
 CC sclerosis, diabetes and allergies. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 212 AA;

Query Match 100.0%; Score 1105; DB 5; Length 212;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTEVDGKVRIG 60  
 Db 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTEVDGKVRIG 60  
 Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDSSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
 Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDSSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
 Qy 121 NKADDEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRKELEGL 180  
 Db 121 NKADDEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRKELEGL 180  
 Qy 181 RMRASNELALAELEEEGKPEGPNANSKTCWC 212  
 Db 181 RMRASNELALAELEEEGKPEGPNANSKTCWC 212

RESULT 2  
 AAEE29096  
 ID AAEE29096 standard; protein; 212 AA.  
 XX AC AAEE29096;  
 XX 24-FEB-2003 (first entry)  
 XX Human Ras-like protein.  
 XX Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
 KW immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
 KW Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
 KW osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
 KW irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
 KW pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
 KW Sjogren's syndrome; infection; transgenic; gene therapy; neutropenic; gout;  
 KW neuroprotective; vasotropic; cytostatic; dermatological; nephrotropic;  
 KW ophthalmological; tranquilizer; cancer; stroke; Grave's disease; AIDS;  
 KW asthma; anaemia; drug screening.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Binding-site 15..22 /note= "ATP/GTP binding site motif A"  
 FT Modified-site 18..23 /note= "N-myristoylation site"  
 FT Modified-site 29..32 /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 92..94 /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 101..109 /note= "Tyrosine kinase phosphorylation site"  
 FT Modified-site 104..107 /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 136..141 /note= "N-myristoylation site"  
 FT Modified-site 205..208 /note= "N-glycosylation site"  
 FT Modified-site 206..208

FT /note= "Protein kinase C phosphorylation site"  
 XX WO200277193-A2.  
 XX 03-OCT-2002.  
 XX 27-MAR-2002; 2002WO-US009328.  
 XX 27-MAR-2001; 2001US-00817198.  
 XX (PEKE ) PE CORP.  
 XX Gan W, Ye J, Di Francesco V, Beasley EM;  
 XX WPI; 2003-018913/01.  
 XX N-PSDB; AAD47168, AAD47619.  
 XX New isolated human Ras-like protein polypeptide, useful for diagnosing,  
 PT treating or preventing inflammation and disorders associated with cell  
 PT proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma  
 PT or stroke.  
 XX  
 PS Claim 1; Page 74; 82pp; English.  
 XX The invention relates to human Ras-like protein and its corresponding  
 CC nucleic acid. The Ras-like protein and DNA is useful in the development  
 CC of human therapeutics and diagnostic compositions. They are useful in the  
 CC diagnosis, prevention and treatment of inflammation and disorders  
 CC associated with cell proliferation and apoptosis, e.g. AIDS and other  
 CC infectious or genetic immunodeficiencies, neurodegenerative disease e.g.  
 CC Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,  
 CC ischaemic injuries e.g. myocardial infarction, stroke or reperfusion  
 CC injury, toxin-induced diseases such as alcohol-induced liver damage or  
 CC cirrhosis, osteoporosis or cancer. They are also used to treat disorders  
 CC associated with inflammation including allergies, atopic dermatitis,  
 CC atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,  
 CC lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,  
 CC autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,  
 CC uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or  
 CC helminthic infections. The antibodies of the invention are useful in  
 CC pharmacogenomic analysis or for tissue typing. The transgenic animals are  
 CC useful for studying the function of a Ras-like protein, and identifying  
 CC and evaluating modulators of its activity. Ras-like protein is used in  
 CC drug screening assays and its DNA is used in gene therapy. The present  
 CC sequence is human Ras-like protein  
 XX  
 SQ Sequence 212 AA;

Query Match 100.0%; Score 1105; DB 6; Length 212;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTEVDGKVRIG 60  
 Db 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTEVDGKVRIG 60  
 Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDSSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
 Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDSSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
 Qy 121 NKADDEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRKELEGL 180  
 Db 121 NKADDEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRKELEGL 180  
 Qy 181 RMRASNELALAELEEEGKPEGPNANSKTCWC 212  
 Db 181 RMRASNELALAELEEEGKPEGPNANSKTCWC 212

RESULT 3  
 AAU17136  
 ID AAU17136 standard; protein; 401 AA.

XX AAU17136;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Novel signal transduction pathway protein, Seq ID 701.  
XX  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX  
XX WO200154733-A1.  
PN  
XX  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001312.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184684P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241321P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465460/50.  
XX N-PSDB; AAS27053.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders.  
XX  
XX Claim 1; SEQ ID NO 701; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anaemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative disorders  
XX (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
XX Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
XX (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
XX respiratory disorders, dermatological disorders, in wound healing,  
XX epithelial cell proliferation, endocrine disorders (e.g. Addison's  
XX disease), reproductive system disorders, gastrointestinal disorder  
XX (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
XX B-cell responsiveness to pathogens, activators of T-cells, to induce  
XX higher affinity antibodies, and as a means to induce tumour proliferation  
XX in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-  
XX AAU17683 represent novel signal transduction pathway protein, amino acid  
XX sequences of the invention

Query Match 100.0%; Score 1105; DB 4; Length 401;  
Best Local Similarity 100.0%; Pred. No. 3.5e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDVGKGTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQ 60  
Db 27 MAKQYDVLFRLLIGDVGKGTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQ 86  
QY 61 IWDTAGQERYQITTKQYRRAQGI FLVYDISERSYQHIMKMWSDVDEYAPGVOKILIG 120  
Db 87 IWDTAGQERYQITTKQYRRAQGI FLVYDISERSYQHIMKMWSDVDEYAPGVOKILIG 146  
QY 121 NKADBEQKQVGRQEQQLAKYGMDFYETSACTNLNLIKESFRTLTELVLQARRKELEGL 180  
Db 147 NKADBEQKQVGRQEQQLAKYGMDFYETSACTNLNLIKESFRTLTELVLQARRKELEGL 206  
QY 181 RMRASNELALAELEBEKPEGPANSSKTCWC 212  
Db 207 RMRASNELALAELEBEKPEGPANSSKTCWC 238

RESULT 4  
ADB93844  
ID ADB93844 standard; protein; 401 AA.  
XX  
XX ADB93844;  
AC  
DT 04-DEC-2003 (first entry)  
XX  
XX Human novel protein #78.  
XX human; autoimmune disease; Parkinson's disease; silicosis;  
KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;  
KW immunosuppressive agent; adjuvant; enhance immune response;  
KW higher affinity antibody induction;  
KW increased serum immunoglobulin concentration.  
XX  
XX Homo sapiens.  
XX  
XX US2002168711-A1.  
XX  
XX 14-NOV-2002.  
XX  
XX 17-JAN-2001; 2001US-00764868.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220363P.  
XX 26-JUL-2000; 2000US-0220364P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 22-AUG-2000; 2000US-0226868P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 21-SEP-2000; 2000US-0234223P.  
XX 21-SEP-2000; 2000US-0234274P.  
XX 25-SEP-2000; 2000US-0234997P.  
XX 27-SEP-2000; 2000US-0235834P.  
XX 29-SEP-2000; 2000US-0236327P.  
XX 29-SEP-2000; 2000US-0236367P.  
XX 29-SEP-2000; 2000US-0236388P.  
XX 29-SEP-2000; 2000US-0236369P.  
XX 29-SEP-2000; 2000US-0236370P.  
XX 02-OCT-2000; 2000US-0236802P.  
XX 02-OCT-2000; 2000US-0237037P.  
XX 02-OCT-2000; 2000US-0237038P.  
XX 02-OCT-2000; 2000US-0237039P.  
XX 02-OCT-2000; 2000US-0237040P.  
XX 13-OCT-2000; 2000US-0239335P.  
XX 20-OCT-2000; 2000US-0240960P.  
XX 20-OCT-2000; 2000US-0241785P.  
XX 20-OCT-2000; 2000US-0241809P.  
XX 01-NOV-2000; 2000US-0244617P.  
XX 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2003-71985/68.  
DR N-PSDB; ADB93221.  
XX  
XX New isolated polypeptide useful for diagnosing and treating  
PT immunosuppressive conditions such as autoimmune disease and Parkinson's  
PT disease.  
XX  
PS Claim 11; SEQ ID NO 701; 345pp; English.  
XX  
XX The invention relates to an isolated polypeptide. The polypeptide is  
CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition in a subject, by determining the presence or  
CC amount of expression of the polypeptide in a biological sample and  
CC diagnosing a pathological condition or a susceptibility to a pathological  
CC condition based on the presence or amount of expression of the  
CC polypeptide. The polypeptide is also useful for identifying a binding  
CC partner to the polypeptide, which involves contacting the polypeptide  
CC with a binding partner and determining whether the binding partner  
CC effects an activity of the polypeptide. The polypeptide or the nucleic  
CC acid encoding the polypeptide is useful for preventing, treating, or  
CC ameliorating a medical condition, which involves administering the  
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
CC is useful for diagnosing a pathological condition or a susceptibility to  
CC a pathological condition in a subject, which involves determining the  
CC presence or absence of a mutation in the nucleic acid, and diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC based on the presence or absence of the mutation. The polypeptide, the  
CC nucleic acid and an antibody to the polypeptide are useful for treating  
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein. Note: The sequence data for this patent did not form part of the  
CC printed specification but was obtained in electronic format direct from  
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20020168711.  
XX  
XX Sequence 401 AA;

Query Match 100.0%; Score 1105; DB 7; Length 401;  
Best Local Similarity 100.0%; Pred. No. 3.5e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPMKTIKVRQ 60  
Db 27 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPMKTIKVRQ 86  
Qy 61 IWDTAGQRYQITTKQYRRAGQIFLVYDTSERSYQIMKWSVDVYAPGVQKILIG 120  
Db 87 IWDTAGQRYQITTKQYRRAGQIFLVYDTSERSYQIMKWSVDVYAPGVQKILIG 146  
Qy 121 NKABEQKRYGREGQQLAKYGMDFYETACTNLNKTESFTRLTELVLQAHKREGL 180  
Db 147 NKABEQKRYGREGQQLAKYGMDFYETACTNLNKTESFTRLTELVLQAHKREGL 206  
Qy 181 RMRASNELALAELEEEKPEGPANSSKTCWC 212  
Db 207 RMRASNELALAELEEEKPEGPANSSKTCWC 238

ID AAB41604 standard; protein; 218 AA.  
XX AAB41604;  
DT 08-FEB-2001 (first entry)  
DE Human ORFX ORF1368 polypeptide sequence SEQ ID NO:2736.  
XX  
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnary; antiparokinon; antiparokinon; antiparokinon; hepatotropic;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.  
XX Homo sapiens.  
OS  
XX WO200058473-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000WO-US008621.  
XX  
XX 31-MAR-1999; 99US-0127607P.  
PR 02-APR-1999; 99US-0127636P.  
PR 05-APR-1999; 99US-0127728P.  
PR 30-MAR-2000; 2000US-00540763.  
XX (CURA-) CURAGEN CORP.  
PI Shimkets RA, Leach M;  
DR WPI; 2000-602362/57.  
DR N-PSDB; AAC75813.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 11; Page 1979-1980; 5507pp; English.  
XX  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
CC antiparokinon; antiparokinon; antiparokinon; hepatotropic; vulnary;  
CC anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
CC immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
CC dermatological; immunosuppressive; antirheumatic; antithyroid; antianemic. The  
CC antiviral; antifungal; antirheumatic; antithyroid; antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
XX Sequence 218 AA;

Best Local Similarity 97.2%; Pred. No. 3.4e-107;  
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
Db 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
Qy 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGEGVKILIG 120  
Db 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGEGVKILIG 120  
Qy 121 NKADEQKQVGRGQGO-----OLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 174  
Db 121 NKADEQKQVGRGQGOQKQPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 180  
Qy 175 KELEGLMRASNELALAELEEEKGPEGPANSSKTCWC 212  
Db 181 KELEGLMRASNELALAELEEEKGPEGPANSSKTCWC 218

RESULT 6  
ID ADE58127 standard; protein; 212 AA.  
AC ADE58127;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein AAA41995, SEQ ID NO 3998.  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; AAA41995.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat  
or human polynucleotides or a polynucleotide which represents a fragment,  
derivative or allelic variation of the nucleic acid sequence. Also  
claimed are a vector comprising the novel polynucleotide, a host cell  
comprising the vector, a method for identifying a nucleotide sequence  
which is differentially regulated in an animal subjected to pain and a  
kit to perform the method, an array, a method for identifying an agent  
that increases or decreases the expression of the polynucleotide sequence  
that is differentially expressed in neuronal tissue of a first animal  
subjected to pain, a method for identifying a compound which regulates  
the expression of a polynucleotide sequence which is differentially  
expressed in an animal subjected to pain, a method for identifying a  
compound that regulates the activity of one or more of the  
polynucleotides, a method for producing a pharmaceutical composition, a  
method for identifying a compound or small molecule that regulates the

activity in an animal of one or more of the polypeptides given in the  
specification, a method for identifying a compound useful in treating  
pain and a pharmaceutical composition comprising the one or more  
polypeptides or their antibodies. The polynucleotide or the compound that  
modulates its activity is useful for preparing a medicament for treating  
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
therapy). The sequence presented is a rat protein (shown in Table 2 of  
the specification) which is differentially expressed during pain. Note:  
The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 212 AA;

Query Match 97.5%; Score 1077; DB 7; Length 212;  
Best Local Similarity 97.6%; Pred. No. 1.3e-105;  
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
Db 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
Qy 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGEGVKILIG 120  
Db 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGEGVKILIG 120  
Qy 121 NKADEQKQVGRGQGOQKQPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAHKELEGL 180  
Db 121 NKADEQKQVGRGQGOQKQPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAHKELEGL 180  
Qy 181 RMRASNELALAELEEEKGPEGPANSSKTCWC 212  
Db 181 RTCASNELALAELEEEKGTEGPANSSKTCWC 212

RESULT 7  
AAU17555  
ID AAU17555 standard; protein; 188 AA.  
XX  
AC AAU17555;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Novel signal transduction pathway protein, Seq ID 1120.  
XX  
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX  
XX WO200154733-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001312.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.



PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225286P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-465460/50.  
N-PSDB; AAS27472.  
Novel polypeptides useful for diagnosing, treating, preventing and/or  
prognosing disorders related to the proteins, including cancers, immune  
disorders and neuronal disorders.  
Claim 1; SEQ ID NO 1120; 880pp; English.  
The invention relates to novel isolated polypeptides (I), and  
polynucleotides (II). (i), (ii) and the antibody to (i) are useful for  
diagnosing, preventing and treating diseases including immune system  
disorders (e.g. congenital and acquired immunodeficiencies, autoimmune

disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
transplant rejections and graft versus host disease, infectious diseases  
(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
other blood-related disorders (sickle cell anaemia), myeloproliferative  
disorders, primary haematopoietic disorders, hyperproliferative disorders  
(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
respiratory disorders, dermatological disorders, in wound healing,  
epithelial cell proliferation, endocrine disorders (e.g. Addison's  
disease), reproductive system disorders, gastrointestinal disorder  
(inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
B-cell responsiveness to pathogens, activators of T-cells, to induce  
higher affinity antibodies, and as a means to induce tumour proliferation  
in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-  
AAU17683 represent novel signal transduction pathway protein, amino acid  
sequences of the invention

Query Match 75.3%; Score 832; DB 4; Length 188;  
Best Local Similarity 97.6%; Pred. No. 1.1e-79;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLLCFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 60  
DB 24 MAKQYDVLFRLLIGDSGVGKTCCLLCFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 83  
QY 61 IWDTAGERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILIG 120  
DB 84 IWDTAGERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILIG 143  
QY 121 NKADEQKRVGRQGOQLAKYEGMDPYETSACTNLIKESFTR 164  
DB 144 NKADEQKRVGRQGOQLAKYEGMDPYETSACTNLIKESFTR 187

RESULT 8  
ADB94263  
ID ADB94263 standard; protein; 188 AA.  
XX  
AC ADB94263;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human novel protein #497.  
XX

human; autoimmune disease; Parkinson's disease; silicosis;  
gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;  
immunosuppressive agent; adjuvant; enhance immune response;  
higher affinity antibody induction;  
increased serum immunoglobulin concentration.

XX Homo sapiens.  
OS  
XX  
XX US2002168711-A1.  
XX  
XX 14-NOV-2002.  
XX  
XX 17-JAN-2001; 2001US-00764868.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218230P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.  
(RUBE/) RUBEN S M.  
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-719985/68.  
N-PSDB; ADB93640.

New isolated polypeptide useful for diagnosing and treating  
immunosuppressive conditions such as autoimmune disease and Parkinson's  
disease.

Claim 11; SEQ ID NO 1120; 345pp; English.

The invention relates to an isolated polypeptide. The polypeptide is  
useful for diagnosing a pathological condition or a susceptibility to a  
pathological condition in a subject, by determining the presence or  
amount of expression of the polypeptide in a biological sample and  
diagnosing a pathological condition or a susceptibility to a pathological  
condition based on the presence or amount of expression of the  
polypeptide. The polypeptide is also useful for identifying a binding  
partner to the polypeptide, which involves contacting the polypeptide  
with a binding partner and determining whether the binding partner  
effects an activity of the polypeptide. The polypeptide or the nucleic  
acid encoding the polypeptide is useful for preventing, treating, or  
ameliorating a medical condition, which involves administering the  
polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
is useful for diagnosing a pathological condition or a susceptibility to  
a pathological condition in a subject, which involves determining the  
presence or absence of a mutation in the nucleic acid, and diagnosing a  
pathological condition or susceptibility to a pathological condition  
based on the presence or absence of the mutation. The polypeptide, the  
nucleic acid and an antibody to the polypeptide are useful for treating

CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein. Note: The sequence data for this patent did not form part of the  
CC printed specification but was obtained in electronic format direct from  
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20020168711.

XX  
SQ Sequence 188 AA;

Query Match 75.3%; Score 832; DB 7; Length 188;  
Best Local Similarity 97.6%; Pred. No. 1.1e-79;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVIQ 60  
Db 24 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVIQ 83  
Qy 61 IWDTAGQERYQITTKQYRRAGQIFLVYDYSERSYQHIMKWSDVDEYAPGVQKILIG 120  
Db 84 IWDTAGQERYQITTKQYRRAGQIFLVYDYSERSYQHIMKWSDVDEYAPGVQXILIG 143  
Qy 121 NKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTR 164  
Db 144 NKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTR 187

RESULT 9  
ABB70670  
ID ABB70670 standard; protein; 204 AA.  
XX  
AC ABB70670;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 38802.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EM;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL14773.

XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 38802; 21pp + Sequence Listing; English.

XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 204 AA;

Query Match 49.5%; Score 546.5; DB 4; Length 204;  
Best Local Similarity 49.8%; Pred. No. 2.3e-49;  
Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;

Qy 1 MAKQ-YDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVI 59  
Db 1 MAKTYDLLFKLLLLIGDSGVGKTCILFRSDDAFTSTFISTIGDFKIKTVELRGKIKL 60  
Qy 60 QIWDTAGQERYQITTKQYRRAGQIFLVYDYSERSYQHIMKWSDVDEYAPGVQKILI 119  
Db 61 QIWDTAGQERYQITTKQYRRAGQIFLVYDYSERSYQHIMKWSDVDEYAPGVQKILI 120  
Qy 120 GNKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTRLTELVLQAHKELEG 179  
Db 121 GNKCDMTDKRVNKRGEATAREHGIRFEMTSKSNINIERAFCELAAILD---KTSG 176  
Qy 180 LRWRASNELALALEEBEGKPEGPANSSKTC 210  
Db 177 RESAENQERVIIIDRRNQEKAP----GYSKCC 203

RESULT 10  
ADE58121  
ID ADE58121 standard; protein; 207 AA.  
XX  
AC ADE58121;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P24407, SEQ ID NO 3992.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P24407.

XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.

XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent

RESULT 11		
ABB11916		
ID	ABB11916	standard; peptide; 213 AA.
XX		
AC	ABB11916;	
XX		
DT	11-JAN-2002	(first entry)
XX		
DE	Human rab8 homologue, SEQ ID NO:2286.	
XX		
KW	Human; cytokine; cell proliferation; cell differentiation; growth factor;	
KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;	
KW	inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;	
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;	
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;	
KW	chronic inflammatory condition; proliferative retinopathy;	
KW	atherosclerosis; coronary heart disease; arterial ischaemia;	
KW	bone disorder; osteoporosis; vascular growth disorder;	
KW	tissue regeneration; wound healing; infection; immune disorder;	
KW	cell culture; drug screening; gene therapy; antiinflammatory;	
KW	antasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;	
KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;	
KW	antifungal; vulnereary; antiulcer.	
XX		
OS	Homo sapiens.	
XX		
FN	WO200157188-A2.	
XX		
PD	09-AUG-2001.	

Db 127 NKCDVNDKRVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGN 186  
Qy 181 RMRASNE 187  
Db 187 SPQGSNQ 193

RESULT 12  
ID ABP41333 standard; protein; 221 AA.  
XX  
AC ABP41333;  
XX  
DT 23-AUG-2002 (first entry)  
XX  
DE Human ovarian antigen HCGMA67, SEQ ID NO:2465.  
XX  
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; reproductive system disorder; infertility;  
KW pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS;  
KW ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
OS Homo sapiens.  
XX  
FN WO200200677-A1.  
XX  
PD 03-JAN-2002.  
XX  
PF 07-JUN-2001; 2001WO-US018569.  
XX  
PR 07-JUN-2000; 2000US-0209467P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Birse CE, Rosen CA;  
XX  
DR WPI; 2002-147878/19.  
DR N-PSDB; ABQ54410.  
XX  
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.  
XX  
PS Claim 11; SEQ ID NO 2465; 2922pp; English.  
XX

The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovarian and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 221 AA;

Query Match 48.9%; Score 540; DB 5; Length 221;  
Best Local Similarity 52.7%; Pred. No. 1.3e-48;  
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 2 AKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTKTIYDGVKVRIOI 61  
Db 16 AKTYDYLFKLLIGDSGVGKTCVLFSEDAFNFTSTIGIDFKRTIYDGVKVRIOI 75  
Qy 62 WDTAGQERYQTITKQYYRRAGQIFLVYDISSERSYQHIMKWSVDVEYAPGVOKILIGN 121  
Db 76 WDTAGQERFRTITAYYRGAMGIMLVYDITNEKSFDRNWRNIEEHASADVEKMILGN 135  
Qy 122 KADEQKRVGRGQGOOLAKYGMDFYETSACTNWNKESPTRLTELVLQHRKELEGLR 181  
Db 136 KCDVNDKRVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGN 195  
Qy 182 MRASNE 187  
Db 196 PQGSNQ 201

RESULT 13

ADE83429  
ID ADE83429 standard; protein; 203 AA.

XX ADE83429;

XX 29-JAN-2004 (first entry)

XX Human Protein P51153, SEQ ID NO 11024.

XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

XX (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P51153.

XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,

```
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI)). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 203 AA;
Query Match 48.1%; Score 532; DB 7; Length 203;
Best Local Similarity 47.6%; Pred. No. 8e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;
QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKVRIO 60
DB 1 MAKAYDHLFKLLIGDSGVGKTCLLIRPAEDNFNNTYISTIGIDFKIRTVDIEGKTKIQ 60
QY 61 IWDTAGOERYOTITKQYVRAAGIFLVYDIDSSERSYOHIMKWVSDVDEYAPGKVKILIG 120
DB 61 VMDTAGOERFKTITTYAYRGAMGILVYDIDSKSFENIQNMWKSIVENASAGVERLLLG 120
QY 121 NKADDEQKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTFLVLOAHRKELEGL 180
DB 121 NKCDMEAKRVKQEQADKLAREHGIRFFETSASKSMNVDEAFSLARDIL-----LKSG 174
QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210
DB 175 GRRSGN-----GNKP--PSTDLTCTC 192
RESULT 14
ID ADE58125 standard; protein; 203 AA.
XX
AC ADE58125;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P51153, SEQ ID NO 3996.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002NO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX
.PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI: 2003-268312/26.
DR GENBANK; P51153.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX injury (CCI)). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 203 AA;
Query Match 48.1%; Score 532; DB 7; Length 203;
Best Local Similarity 47.6%; Pred. No. 8e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;
QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKVRIO 60
DB 1 MAKAYDHLFKLLIGDSGVGKTCLLIRPAEDNFNNTYISTIGIDFKIRTVDIEGKTKIQ 60
QY 61 IWDTAGOERYOTITKQYVRAAGIFLVYDIDSSERSYOHIMKWVSDVDEYAPGKVKILIG 120
DB 61 VMDTAGOERFKTITTYAYRGAMGILVYDIDSKSFENIQNMWKSIVENASAGVERLLLG 120
QY 121 NKADDEQKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTFLVLOAHRKELEGL 180
DB 121 NKCDMEAKRVKQEQADKLAREHGIRFFETSASKSMNVDEAFSLARDIL-----LKSG 174
QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210
DB 175 GRRSGN-----GNKP--PSTDLTCTC 192
RESULT 15
ID AAB58196 standard; protein; 246 AA.
XX
XX AAB58196;
XX
XX 14-MAR-2001 (first entry)
DT
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:05:27 ; Search time 23 Seconds  
(without alignments)  
475.857 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLVRLLLIGDSVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*
- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/iaa/PCFUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	545	49.3	207	2	US-08-824-873-4
2	545	49.3	207	3	US-09-198-184-4
3	527.5	47.7	205	2	US-08-531-525-25
4	527.5	47.7	205	2	US-08-718-270A-25
5	518.5	46.9	198	2	US-08-531-525-51
6	518.5	46.9	198	2	US-08-718-270A-51
7	515.5	46.7	203	4	US-09-255-920A-12
8	503.5	45.6	207	2	US-08-531-525-35
9	503.5	45.6	207	2	US-08-718-270A-35
10	494.5	44.8	215	2	US-08-531-525-10
11	494.5	44.8	215	2	US-08-718-270A-10
12	488.5	44.2	194	2	US-08-531-525-34
13	488.5	44.2	194	2	US-08-718-270A-34
14	478.5	43.3	201	4	US-09-154-602-3
15	478.5	43.3	201	4	US-08-916-901-3
16	469	42.4	201	2	US-08-531-525-10
17	469	42.4	201	2	US-08-718-270A-10
18	463	41.9	202	2	US-08-531-525-14
19	463	41.9	202	2	US-08-718-270A-14
20	453.5	41.0	201	2	US-08-531-525-13
21	453.5	41.0	201	2	US-08-718-270A-13
22	452.5	41.0	190	2	US-08-824-873-3
23	452.5	41.0	190	3	US-09-198-184-3
24	452.5	41.0	190	4	US-09-255-920A-7
25	443	40.1	218	2	US-08-531-525-19
26	443	40.1	218	2	US-08-718-270A-19
27	430.5	39.0	217	4	US-09-255-920A-15

28	430	38.9	202	4	US-09-255-920A-11	Sequence 11, Appl
29	430	38.9	216	4	US-09-255-920A-14	Sequence 14, Appl
30	416	37.6	191	4	US-09-075-454-3	Sequence 3, Appl
31	412.5	37.3	212	4	US-09-399-913-67	Sequence 67, Appl
32	409.5	37.1	190	2	US-08-824-873-1	Sequence 1, Appl
33	409.5	37.1	190	3	US-09-198-184-1	Sequence 1, Appl
34	409.5	37.1	212	2	US-08-531-525-18	Sequence 18, Appl
35	409.5	37.1	212	2	US-08-718-270A-18	Sequence 18, Appl
36	409.5	37.1	214	2	US-08-531-525-52	Sequence 52, Appl
37	409.5	37.1	214	2	US-08-718-270A-52	Sequence 52, Appl
38	396.5	35.9	203	2	US-08-766-551-8	Sequence 8, Appl
39	396	35.8	192	4	US-09-255-920A-5	Sequence 5, Appl
40	388.5	35.2	213	4	US-09-976-594-1097	Sequence 1097, Ap
41	387	35.0	210	2	US-08-531-525-16	Sequence 16, Appl
42	387	35.0	210	2	US-08-718-270A-16	Sequence 16, Appl
43	373	33.8	213	2	US-08-773-423-8	Sequence 8, Appl
44	373	33.8	217	2	US-08-773-423-3	Sequence 3, Appl
45	372	33.7	208	2	US-08-531-525-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-08-824-873-4  
; Sequence 4, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,873  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 207 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 234746  
; US-08-824-873-4

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.4e-52;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTIIEVDGKIRIQ 60  
Db 1 MAKYDYLFRLLIGDSGVGKTCVLFPSDFANFTISIGIDFKIRTIELDGKRIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKMWSDVDEYAPGVQKILIG 120  
Db 61 IWDTAGQERFTITTAYYRGAMGIMLVYDITNEKSPDNIRNWIENAFPTLARDIKAKMDKKEGN 120  
QY 121 NKADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGL 180  
Db 121 NKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKEGN 180  
QY 181 RMRASNE 187  
Db 181 SPQGSNQ 187

RESULT 2  
US-09-198-184-4  
; Sequence 4, Application US/09198184  
; Patent No. 6010859  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/198,184  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,873  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 207 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 234746  
US-09-198-184-4

Query Match 49.3%; Score 545; DB 3; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.4e-52;  
Matches .99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTIIEVDGKIRIQ 60  
Db 1 MAKYDYLFRLLIGDSGVGKTCVLFPSDFANFTISIGIDFKIRTIELDGKRIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKMWSDVDEYAPGVQKILIG 120

Db 61 IWDTAGQERFTITTAYYRGAMGIMLVYDITNEKSPDNIRNWIENAFPTLARDIKAKMDKKEGN 120  
QY 121 NKADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGL 180  
Db 121 NKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKEGN 180  
QY 181 RMRASNE 187  
Db 181 SPQGSNQ 187

RESULT 3  
US-08-531-525-25  
; Sequence 25, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; TITLE OF INVENTION: of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/531,525  
; FILING DATE: 21-SEP-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feiber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 37-94  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 205 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Canis familiaris  
US-08-531-525-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;  
Best Local Similarity 52.4%; Pred. No. 2.1e-50;  
Matches .97; Conservative 45; Mismatches 42; Indels 1; Gaps 1;  
QY 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTIIEVDGKIRIQ 62  
Db 2 KYDYLFKLLIGDSGVGKTCVLFPSDFANFTISIGIDFKIRTIELDGKRIKQ 61  
QY 63 DTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKMWSDVDEYAPGVQKILGNK 122  
Db 62 DTAGQERFTITTAYY-RAMGIMLVYDITNEKSPDNIRNWIENAFPTLARDIKAKMDKKEGN 120  
QY 123 ADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGLRM 182

Db 121 CDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 180

Qy 183 RASNE 187

Db 181 QGSNQ 185

## RESULT 4

US-08-718-270A-25  
; Sequence 25, Application US/08718270A  
; Patent No. 5910478

## GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5910478le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: US

ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/718,270A

FILING DATE: 20-SEP-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/531,525

FILING DATE: 21-SEP-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/004,091

FILING DATE: 21-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Ferber, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 78-95

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 205 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Canis familiaris

US-08-718-270A-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;

Best Local Similarity 52.4%; Pred. No. 2.1e-50;

Matches 97; Conservative 45; Mismatches 4; Indels 1; Gaps 1;

Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIENVGDKVRIQ 62

Db 2 KTYDVLFRLLIGDSGVGKTCVLFPSDANSTFISIGIDFKRTIELDGKRIKLQIW 61

Qy 63 DTAGERYQTITKQYRRAQGIPLVYDISERSYOHIMKWSDVDEYAPGVQKILGNK 122

Db 62 DTAGERYQTITAYY-RAMGIMLVYDITNEKSFNIRNWRINTEBHASADVERKILGNK 120

Qy 123 ADEBQKQVGRQEQQLAKYGMDFYFETSACTNINIKESFTRLTFLVLAQHRKELEGURM 182

Db 121 CDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 180

Qy 183 RASNE 187

Db 181 QGSNQ 185

## RESULT 5

US-08-531-525-51

; Sequence 51, Application US/08531525

; Patent No. 5840683

## GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5840683le, John F.  
APPLICANT: Abajian, Henry B.

APPLICANT: Kende, Andrew S.

TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action

TITLE OF INVENTION: of P21 Ras

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee and Winner, P.C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: US

ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/531,525

FILING DATE: 21-SEP-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ferber, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 37-94

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 198 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Canis familiaris

US-08-531-525-51

Query Match 46.9%; Score 518.5; DB 2; Length 198;

Best Local Similarity 55.0%; Pred. No. 2e-49;

Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIENVGDKVRIQ 60

Db 1 MKKYDVLFRLLIGDSGVGKTCVLFPSDDAFNTTIS-IGIDFKITVLEQKKIKLQ 59

Qy 61 IWDTAGERYQTITKQYRRAQGIPLVYDISERSYOHIMKWSDVDEYAPGVQKILIG 120

Db 60 IWDTAGERPHITTSYYRGAMGIMLVYDITNGKSFENISKWLARNIDSHANEDVERMLLG 119

Qy 121 NKADDEKQVGRQGGQOLAKYGMDFYFETSACTNINIKESFTRLTFLVLQ 171

Db 120 NKCDMDKVVYPKGEQIAREHGIRFRETSAKVNNINIEKAFLTLAEDILR 170

RESULT 6  
US-08-718-270A-51  
; Sequence 51, Application US/08718270A  
; Patent No. 5910478  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5910478le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptidomimetics Inhibiting  
; TITLE OF INVENTION: The Oncogenic Action of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/718,270A  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/531,525  
; FILING DATE: 21-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,091  
; FILING DATE: 21-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 78-95  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Canis familiaris  
US-08-718-270A-51  
Query Match 46.9%; Score 518.5; DB 2; Length 198;  
Best Local Similarity 55.0%; Pred. No. 2e-49;  
Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRIQ 60  
Db 1 MKKTYDLLFKLLIGDSGVGKTCVLFRESDDAFNTTIS-IGIDFKIKTVLQKKIKIQ 59  
Qy 61 IWDTAGERYQITIKQYRRAGQIFLVYDISSERSYQHKMKWSDVDEYAPGVQKILIG 120  
Db 60 IWDTAGQERFHTITTSYRAGMGLVYDITNGKSFENISKWLRNIDEHANEDVERMLLG 119  
Qy 121 NKADEEQKQVREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQ 171  
Db 120 NKCDMDKRVVPGKGEGIAREHGIRFFETSAKVNINIEKAFLTLAEDILR 170  
RESULT 7

US-09-255-920A-12  
; Sequence 12, Application US/09255920A  
; Patent No. 6623980  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Joseph  
; APPLICANT: Lorens, James  
; APPLICANT: Anderson, David  
; APPLICANT: Luo, Ying  
; APPLICANT: Huang, Betty  
; APPLICANT: Shen, Mary  
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS  
; FILE REFERENCE: A65905-1/DJB/RMS  
; CURRENT APPLICATION NUMBER: US/09/255,920A  
; CURRENT FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: 60/075,534  
; PRIOR FILING DATE: 1998-02-23  
; PRIOR APPLICATION NUMBER: 60/086,650  
; PRIOR FILING DATE: 1998-05-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Mouse  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (34)  
; OTHER INFORMATION: The xaa at position 34 represents an unknown amino  
; OTHER INFORMATION: acid.  
US-09-255-920A-12  
Query Match 46.7%; Score 515.5; DB 4; Length 203;  
Best Local Similarity 56.0%; Pred. No. 4.4e-49;  
Matches 93; Conservative 41; Mismatches 31; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEF-HSHSHISTIGVDFKMTIEVDGKVR 59  
Db 1 MAKTYDYLFRLLIGDSGVGKTCVLFRESDDAFNSTFISTIGIDFKIRTIELDKRIKL 60  
Qy 60 QIWDTAGERYQITIKQYRRAGQIFLVYDISSERSYQHKMKWSDVDEYAPGVQKILI 119  
Db 61 QIWDTAGQERFHTITTSYRAGMGLVYDITNEKSFNIRNIRNIEEHASADVEKML 120  
Qy 120 GNKADBEQKQVREGQOOLAKKEYGMDFYETSACTNLNIKESFTRL 165  
Db 121 GNKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFTTL 166  
RESULT 8  
US-08-531-525-35  
; Sequence 35, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; TITLE OF INVENTION: of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Discopyge ommata  
US-08-531-525-35

Query Match 45.6%; Score 503.5; DB 2; Length 207;  
Best Local Similarity 46.4%; Pred. No. 9.8e-48;  
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;

Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDYDFKMKTEVDGKVRQIW 62  
Db 2 KTYDYLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDYDFKMKTEVDGKVRQIW 61  
Qy 63 DTAGQERYQTITKYYRRAQGIPLVYDISERSYOHIMKMWSDVDEVAPEGVQKILGNK 122  
Db 62 DTAGQERFRTIT-AYYRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
Qy 123 ADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFTRLTLVLQAHKLEGLRM 182  
Db 120 CDNNEKQVSKERGEKLAIDYGIKFLTSKASSINVEEAFITLARDIMTKLNKMM----- 174  
Qy 183 RASNELALAE-LEEEGKPEGPANSSK 208  
Db 175 ---NENSLQEAVDKLGKPPKPSQKK 198

RESULT 9  
US-08-718-270A-35  
Sequence 35, Application US/08/18270A  
Patent No. 5910478  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5910478le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-718-270A-35

Query Match 45.6%; Score 503.5; DB 2; Length 207;  
Best Local Similarity 46.4%; Pred. No. 9.8e-48;  
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;

Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDYDFKMKTEVDGKVRQIW 62  
Db 2 KTYDYLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDYDFKMKTEVDGKVRQIW 61  
Qy 63 DTAGQERYQTITKYYRRAQGIPLVYDISERSYOHIMKMWSDVDEVAPEGVQKILGNK 122  
Db 62 DTAGQERFRTIT-AYYRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
Qy 123 ADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFTRLTLVLQAHKLEGLRM 182  
Db 120 CDNNEKQVSKERGEKLAIDYGIKFLTSKASSINVEEAFITLARDIMTKLNKMM----- 174  
Qy 183 RASNELALAE-LEEEGKPEGPANSSK 208  
Db 175 ---NENSLQEAVDKLGKPPKPSQKK 198

RESULT 10  
US-08-531-525-10  
Sequence 10, Application US/08531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5840683le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525

```
;
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Arabidopsis thaliana
;
US-08-531-525-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 1e-46;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLIGDSGVGKSCLLRSDFSTFTTIGIDFKITIEDGKRKILQIWD 71
Qy 65 AGQERYOTITKQYVRAQGIPLVYDISERSYQHIMKWSVDVYAPGVQKILIGNKAD 124
Db 72 AQGER-RTITAYYRGAMGILLVYDVTDESFNIRNIRWIRNIEQHASDNVNVKILVGNKAD 130
Qy 125 -EEQKRVGREGQQOLAKEYGMDPYETSACTNINIKESFTLTLVLQAHKKELEGLRMR 183
Db 131 MDESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEVFFSIG-----RDIKQR 179
Qy 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAGQATQKSAC 212

RESULT 11
US-08-718-270A-10
; Sequence 10, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 59104781e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of p21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
```

```
;
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Arabidopsis thaliana
;
US-08-718-270A-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 1e-46;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLIGDSGVGKSCLLRSDFSTFTTIGIDFKITIEDGKRKILQIWD 71
Qy 65 AGQERYOTITKQYVRAQGIPLVYDISERSYQHIMKWSVDVYAPGVQKILIGNKAD 124
Db 72 AQGER-RTITAYYRGAMGILLVYDVTDESFNIRNIRWIRNIEQHASDNVNVKILVGNKAD 130
Qy 125 -EEQKRVGREGQQOLAKEYGMDPYETSACTNINIKESFTLTLVLQAHKKELEGLRMR 183
Db 131 MDESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEVFFSIG-----RDIKQR 179
Qy 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAGQATQKSAC 212

RESULT 12
US-08-531-525-34
; Sequence 34, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 58406831e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of p21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
```

ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-531-525-34

Query Match 44.2%; Score 488.5; DB 2; Length 194;  
Best Local Similarity 55.8%; Pred. No. 4.1e-46;  
Matches 96; Conservative 41; Mismatches 30; Indels 5; Gaps 4;

Qy 1 MAKQ-YDVLFRLLIGSGVGTCLCRFTDNEPHSSHISTIGVDKMTIEVDGIKVI 59  
Db 1 MAKTYDLLFKLLIGSGVGTCLVFRSDDAFTTIFSTIGIDFKIKTVLHGKKIKL 60

Qy 60 QIWDTAQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKWSDVDVEAPEGVQKILI 119  
Db 61 QIWDTAQERPHIT-SYIRGAMGIMLVYDITNAKSFENISKWLNRNIDEHANEDVERMLL 119

Qy 120 GNKADEQKRGVQEQQLAKYGMDFYTSACTNLNLIKESFRTLVELVQ 171  
Db 120 GNK-DMEDKRVLSKQ--IAEHAIRFFETSAKANINIEKAFILTAEIDIQ 168

RESULT 13  
US-08-718-270A-34  
Sequence 34, Application US/08718270A  
Patent No. 5910478  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5910478le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.

REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-718-270A-34

Query Match 44.2%; Score 488.5; DB 2; Length 194;  
Best Local Similarity 55.8%; Pred. No. 4.1e-46;  
Matches 96; Conservative 41; Mismatches 30; Indels 5; Gaps 4;

Qy 1 MAKQ-YDVLFRLLIGSGVGTCLCRFTDNEPHSSHISTIGVDKMTIEVDGIKVI 59  
Db 1 MAKTYDLLFKLLIGSGVGTCLVFRSDDAFTTIFSTIGIDFKIKTVLHGKKIKL 60

Qy 60 QIWDTAQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKWSDVDVEAPEGVQKILI 119  
Db 61 QIWDTAQERPHIT-SYIRGAMGIMLVYDITNAKSFENISKWLNRNIDEHANEDVERMLL 119

Qy 120 GNKADEQKRGVQEQQLAKYGMDFYTSACTNLNLIKESFRTLVELVQ 171  
Db 120 GNK-DMEDKRVLSKQ--IAEHAIRFFETSAKANINIEKAFILTAEIDIQ 168

RESULT 14  
US-08-916-901-3  
Sequence 3, Application US/08916901  
Patent No. 5892012  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids

CLONE: 2514506  
US-09-154-602-3  
Query Match 43.3%; Score 478.5; DB 4; Length 201;  
Best Local Similarity 45.0%; Pred. No. 5.5e-45;  
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDPMKMTIEVDGKVRQ 60  
DB 1 MNPEYDYLFRLLLLIGDSGVGKSCLLLRFPADDTYTESYISTIGVDPKRTIELDGTIKLQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIIFLYVDIISSESYQHIMKWSVDVDEYAPGVQKILIG 120  
DB 61 IWDTAGQERFRITTSYRGAHGIIIVYDVTQESYANVKQWLQEI DRYASENVNKLVG 120  
QY 121 NKADEEQKRGVREGQQOLAKEYGMDFYETSACTNLNIKESFTLTELVLQAHKLEGL 180  
DB 121 NKSDLTKKVVNDTAKFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKKMGMP 175  
QY 181 RMRASNELALAELEEEBEGKPEG 202  
DB 176 GAASGGERPNLKIDSTPVKPG 197  
Search completed: March 15, 2004, 11:09:10  
Job time : 23 secs

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVRTUT04  
CLONE: 2514506  
US-08-916-901-3  
Query Match 43.3%; Score 478.5; DB 2; Length 201;  
Best Local Similarity 45.0%; Pred. No. 5.5e-45;  
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDPMKMTIEVDGKVRQ 60  
DB 1 MNPEYDYLFRLLLLIGDSGVGKSCLLLRFPADDTYTESYISTIGVDPKRTIELDGTIKLQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIIFLYVDIISSESYQHIMKWSVDVDEYAPGVQKILIG 120  
DB 61 IWDTAGQERFRITTSYRGAHGIIIVYDVTQESYANVKQWLQEI DRYASENVNKLVG 120  
QY 121 NKADEEQKRGVREGQQOLAKEYGMDFYETSACTNLNIKESFTLTELVLQAHKLEGL 180  
DB 121 NKSDLTKKVVNDTAKFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKKMGMP 175  
QY 181 RMRASNELALAELEEEBEGKPEG 202  
DB 176 GAASGGERPNLKIDSTPVKPG 197

RESULT 15  
US-09-154-602-3  
Sequence 3, Application US/09154602  
Patent No. 6300472  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,602  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/916,901  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVRTUT04



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:08:07 ; Search time 34 Seconds  
(without alignments)  
1316.603 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLVFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	9	US-09-817-198A-2
2	1105	100.0	401	9	US-09-764-868-701
3	1092	98.8	218	9	US-09-817-198A-5
4	1077	97.5	212	9	US-09-817-198A-4
5	832	75.3	188	9	US-09-764-868-1120
6	559	50.6	224	15	US-10-369-493-4997
7	540	48.9	221	15	US-10-264-049-2465
8	532	48.1	246	9	US-09-925-302-534
9	530.5	48.0	218	9	US-09-925-300-1571
10	528.5	47.8	207	9	US-09-794-257-8
11	523	47.3	206	14	US-10-128-714-3241
12	523	47.3	206	14	US-10-128-714-8600
13	514	46.5	201	14	US-10-179-766-6
14	514	46.5	201	15	US-10-369-493-5076
15	512	46.3	190	9	US-09-822-860-5

16	497	45.0	162	9	US-09-834-765-766	Sequence 766, Appli
17	478.5	43.3	201	9	US-09-967-736-3	Sequence 3, Appli
18	478.5	43.3	201	15	US-10-291-172-193	Sequence 193, Appl
19	478.5	43.3	224	14	US-10-102-806-466	Sequence 466, Appl
20	473.5	42.9	205	15	US-10-369-493-6261	Sequence 6261, Ap
21	473	42.8	210	14	US-10-032-585-7300	Sequence 7300, Ap
22	470.5	42.6	215	15	US-10-369-493-21876	Sequence 21876, A
23	469.5	42.5	225	14	US-10-128-714-8241	Sequence 8241, Ap
24	469	42.4	201	9	US-09-967-736-8	Sequence 8, Appli
25	467	42.3	222	9	US-09-820-003A-4	Sequence 4, Appli
26	462	41.8	207	14	US-10-032-585-7303	Sequence 7303, Ap
27	456.5	41.3	206	15	US-10-369-493-1787	Sequence 1787, Ap
28	448.5	40.6	216	15	US-10-369-493-5077	Sequence 5077, Ap
29	439	39.7	220	15	US-10-116-275-150	Sequence 150, Appl
30	434.5	39.3	198	9	US-09-794-257-16	Sequence 16, Appli
31	434.5	39.3	198	9	US-09-945-173-5	Sequence 5, Appli
32	434.5	39.3	198	9	US-09-972-529-4	Sequence 4, Appli
33	432	39.1	219	15	US-10-116-275-138	Sequence 138, Appl
34	430	38.9	223	9	US-09-817-199A-4	Sequence 4, Appli
35	429	38.8	223	9	US-09-817-199A-2	Sequence 2, Appli
36	429	38.6	223	15	US-10-108-260A-4746	Sequence 4746, Ap
37	427	38.6	226	9	US-09-764-868-684	Sequence 684, Appl
38	426	38.6	222	9	US-09-764-868-1106	Sequence 1106, Ap
39	421	38.1	214	15	US-10-369-493-5001	Sequence 5001, Ap
40	416	37.6	191	9	US-09-794-257-14	Sequence 14, Appli
41	416	37.6	191	13	US-10-051-986-3	Sequence 3, Appli
42	415.5	37.6	191	14	US-10-258-107-1	Sequence 1, Appli
43	413	37.4	223	15	US-10-369-493-1738	Sequence 1738, Ap
44	412.5	37.3	212	9	US-09-350-874-67	Sequence 67, Appli
45	412.5	37.3	212	14	US-10-106-989-67	Sequence 67, Appli

ALIGNMENTS

RESULT 1  
US-09-817-198A-2  
; Sequence 2, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817,198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-817-198A-2

Query Match	100.0%;	Score 1105;	DB 9;	Length 212;
Best Local Similarity	100.0%;	Pred. No. 2.2e-106;		
Matches 212;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAKQYDLVFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIIVDGIKVRIQ	60	
Db	1	MAKQYDLVFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIIVDGIKVRIQ	60	
Qy	61	IWDTAGERYQTITIKQYRRAQGIFLVYDIDSSERSYQHIMKWSVDVDAEYAEQVKILIG	120	
Db	61	IWDTAGERYQTITIKQYRRAQGIFLVYDIDSSERSYQHIMKWSVDVDAEYAEQVKILIG	120	
Qy	121	NKADEEOKRVGREGGQOLAKKEYGMDVFYETSACNLNLIKESFTRITELVLQAHKKEGL	180	
Db	121	NKADEEOKRVGREGGQOLAKKEYGMDVFYETSACNLNLIKESFTRITELVLQAHKKEGL	180	
Qy	181	RMRASNELALAELEEEGKPEGPANSSKTCWC	212	
Db	181	RMRASNELALAELEEEGKPEGPANSSKTCWC	212	

```
Db      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

RESULT 2
US-09-764-868-701
; Sequence 701, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 701
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-701

Query Match      100.0%; Score 1105; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.4e-106;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 60
Db      27 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 86

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 120
Db      87 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 146

QY      121 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 180
Db      147 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 206

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      207 RMRASNELALAELEEEGKPEGPANSSKTCWC 238

RESULT 3
US-09-817-198A-5
; Sequence 5, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
US-09-817-198A-5

Query Match      98.8%; Score 1092; DB 9; Length 218;
Best Local Similarity 97.2%; Pred. No. 5.2e-105;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 60
Db      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 60

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 120

Db      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 120

Db      121 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 180
Db      121 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 180

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      181 RTCASNELALAELEEDGKTEGPGANSSKTCWC 212

RESULT 4
US-09-817-198A-4
; Sequence 4, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-817-198A-4

Query Match      97.5%; Score 1077; DB 9; Length 212;
Best Local Similarity 97.6%; Pred. No. 1.8e-103;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 60
Db      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGIKVRIQ 60

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 120
Db      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYAPEGVQKILIG 120

QY      121 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 180
Db      121 NKADEEQKRVGREGQQLAKYGMDFYETSACTNINIKESFTRLTELVLQAHKLEGL 180

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      181 RTCASNELALAELEEDGKTEGPGANSSKTCWC 212

RESULT 5
US-09-764-868-1120
; Sequence 1120, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

NAME/KEY: SITE  
LOCATION: (151)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (161)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (164)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (188)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-1120

Query Match 75.3%; Score 832; DB 9; Length 188;  
Best Local Similarity 97.6%; Pred. No. 3.9e-78;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQ 60  
Db 24 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQ 83  
Qy 61 IWDTAGQERYOTITKQYRRAGGIFLVYDYSERSYQHIMKWSVDVEYAPGVQKILIG 120  
Db 84 IWDTAGQERYOTITKQYRRAGGIFLVYDYSERSYQHIMKWSVDVEYAPGVQKILIG 143  
Qy 121 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTR 164  
Db 144 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTR 187

RESULT 6  
US-10-369-493-4997  
Sequence 4997, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 4997  
LENGTH: 224  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-4997

Query Match 50.6%; Score 559; DB 15; Length 224;  
Best Local Similarity 55.3%; Pred. No. 1.1e-49;  
Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDFNSFSTIGIDPKIRTIELDGKKIKLQ 60  
Qy 61 IWDTAGQERYOTITKQYRRAGGIFLVYDYSERSYQHIMKWSVDVEYAPGVQKILIG 120  
Db 61 IWDTAGQERYOTITKQYRRAGGIFLVYDITNERSFENIKWIRNIESHAASDVERMTIG 120  
Qy 121 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGL 180  
Db 121 NKCDIEERVRDRGEQALIEYGTKEFLETSKANLNIDEAFFTLARDI-----KSKMEQN 176  
Qy 181 RMRASNEAL 190  
|||:|:

Db 177 EMRAATGAAI 186  
RESULT 7  
US-10-264-049-2465  
Sequence 2465, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133P1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: Patentin Ver. 3.1  
SEQ ID NO 2465  
LENGTH: 221  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (2)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (6)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (7)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-2465

Query Match 48.9%; Score 540; DB 15; Length 221;  
Best Local Similarity 52.7%; Pred. No. 9.7e-48;  
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
Qy 2 AKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQ 61  
Db 16 AKTYDYLKLLIGDSGVGKTCVLFPSDFNSFSTIGIDPKIRTIELDGKKIKLQ 75  
Qy 62 WDTAGQERYOTITKQYRRAGGIFLVYDYSERSYQHIMKWSVDVEYAPGVQKILIG 121  
Db 76 WDTAGQERYOTITKQYRRAGGIFLVYDITNERSFENIKWIRNIESHAASDVERMTIG 135  
Qy 122 KADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGL 181  
Db 136 KCDVNDKRVSKERGEKALDYGIKFMETSAKAINVENAFFTLARDIKAKMDKKLEGS 195  
Qy 182 MRASNE 187  
Db 196 PQGSNQ 201  
|||:

RESULT 8  
US-09-925-302-534  
Sequence 534, Application US/09925302  
Patent No. US20020044941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925,302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896





Db 1 MARRPYDMLFKLLIGDSGVGKTCILYRFSDDAFNTTFTSTIGIDFKIKTIELKGKIKL 60  
Qy 60 QIWDTAGQERYQITTKOYRRAQGIPLVYDIISSERSYOHIMKWVSDVDEYAPGVQKILI 119  
Db 61 QIWDTAGQERFHTTTSYRGAGIMLVYDITNAKSPDNIAKWLNRNIDHASEDVVVMIL 120  
Qy 120 GNKADEQKQVGRQOQQLAKYGMDFYTSACTNINIKESFTRLTELVLQHRKELEG 179  
Db 121 GNKCDMSDRRVSRERGEKIAQDHGISHFETSAKLVHVDYAFYDLAEAIL----- 171  
Qy 180 LRMRASNELALAELEEEGKPEGPANSKTCWC 212  
Db 172 AKMPDSTD-----EQSRDTVNPVQPRQSSGGC 200

## RESULT 15

US-09-822-860-5  
; Sequence 5, Application US/09822860  
; Patent No. US20020146795A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHU, Shiaoqing et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001214  
; CURRENT APPLICATION NUMBER: US/09/822,860  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Discopyge ommata  
US-09-822-860-5

Query Match 46.3%; Score 512; DB 9; Length 190;  
Best Local Similarity 47.7%; Pred. No. 6.3e-45;  
Matches 93; Conservative 52; Mismatches 42; Indels 8; Gaps 1;  
Qy 6 DVLFRLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMTIEVDGKVRQIWDTA 65  
Db 1 DYLFRLLIGDSGVGKTCILFRSEDAFNTTFTSTIGIDFKIRTVELDGKKIKIWDTA 60  
Qy 66 GQERYQITTKOYRRAQGIPLVYDIISSERSYOHIMKWVSDVDEYAPGVQKILIGNKADE 125  
Db 61 GQERFRTITTYRGAGIMKVVYDITNEKSPDNIAKWLNRNIDHASEDVVVMILGNKCDM 120  
Qy 126 EQKQVGRQOQQLAKYGMDFYTSACTNINIKESFTRLTELVLQHRKELEGRLRAS 185  
Db 121 NEKQVSKERGEKLAIDYGIKFLKTSKSSINVEEAFITLARDIMTKLNKKM----- 172  
Qy 186 NELALAELEEEGKP 200  
Db 173 NENSLQEAVDKIKSP 187

Search completed: March 15, 2004, 11:13:37  
Job time : 35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:03:18 ; Search time 2744 Seconds  
(without alignments)  
3348.660 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLFRLLIGDSGVG.....LEEEKGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09817198/runat\_15032004\_101746\_20542/app\_query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN 1 1 5265 @runat\_15032004\_101746\_20542 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	1054	6	AX399903 Sequence
2	1090	98.6	3139	10	BC027769 Mus muscu
3	1077	97.5	945	10	M83679 Sprague-Daw
4	1048.5	94.9	3326	9	BC040679 Homo sapi
5	938	84.9	2348	5	BC063736 Xenopus l
6	798	72.2	2560	10	BC013790 Mus muscu
7	797.5	72.2	2210	9	BX640825 Homo sapi
8	563.5	51.0	1726	5	BC053195 Danio rer
9	546.5	49.5	1613	3	AB006189 Drosophil
10	546.5	49.5	2349	3	AY060425 Drosophil
11	545	49.3	624	9	AF498943 Homo sapi
12	545	49.3	624	9	BT007184 Homo sapi
13	545	49.3	624	12	BT008275 Synthetic
14	545	49.3	660	9	X56741 H.sapiens m
15	545	49.3	760	4	CFRAB8
16	545	49.3	2048	9	BC002977 Homo sapi
17	545	49.3	2818	9	AK025165 Homo sapi
18	541	49.0	1337	10	BC019990 Mus muscu
19	539.5	48.8	765	5	M83391 Discopoge o
20	534.5	48.4	1211	9	BC000799 Homo sapi
21	534.5	48.4	1238	9	X75593 H.sapiens m
22	534.5	48.4	1356	5	BC060015 Xenopus l
23	532	48.1	612	9	AF498948 Homo sapi
24	531.5	48.1	840	5	AF498948 Homo sapi
25	530.5	48.0	603	9	AF297660 Homo sapi
26	530.5	48.0	603	9	AF498945 Homo sapi
27	530.5	48.0	888	4	CFRAB10
28	530.5	48.0	897	10	AF035646 Mus muscu
29	530.5	48.0	1029	9	BC000896 Homo sapi
30	530.5	48.0	1071	9	AF086917 Homo sapi
31	530.5	48.0	3164	9	AF106681 Homo sapi
32	530.5	48.0	3185	10	BC056374 Mus muscu
33	530.5	48.0	3533	6	AX882713 Sequence
34	530.5	48.0	3533	6	BD159881 Primer fo
35	530.5	48.0	3533	9	AK023223 Homo sapi
36	530.5	48.0	3615	10	BC052735 Mus muscu
37	529.5	47.9	682	8	BT001952 Arabidops
38	529.5	47.9	708	8	BT002186 Arabidops
39	529.5	47.9	954	8	ATHAR3
40	529.5	47.9	969	8	AY042795 Arabidops
41	529.5	47.9	1143	8	AY035132 Arabidops
42	528.5	47.8	624	6	AX236078 Sequence
43	528.5	47.8	740	10	RNU53475 Rattus norv
44	528.5	47.8	759	10	AF525280 Rattus no
45	528.5	47.8	1128	9	BC020654 Homo sapi

ALIGNMENTS

RESULT 1









US-09-817-198c-2 (1-212) x BC040679 (1-3326)	
Qy	1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db	5 ATGGCGAAGCAGTACGATGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGC 64
Qy	21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db	65 AAGACCTGCCTGCTGTCGGCTTCCACGACAAACGAGTTCCTCCTCGCACATCTCCACC 124
Qy	41 IleglyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db	125 ATCCGCTGTGATCTTAAGATGAACACCATAGAGTAGACGCGCATCAAGTCCGATACAG 184
Qy	61 IleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db	185 ATCTGGACACATCGAGGCGAGAGATACAGACCATCACAAGCAGTACTATCGCGG 244
Qy	81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db	245 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCAGCGCTCTTACCAGCACATCATG 304
Qy	101 LysTyrValSerAspValAspGlu----- 108
Db	305 AAGTGGGTCACTGACCTGGATGA-GGTAGGAGATGCCACCTCACTGCCGGGTGTGGAGA 363
Qy	108 ----- 108
Db	364 GGTGCTCACCAGGGAAGGCAAGCGGCGGCGAGTGGGAGGCAATGCTTCAGGAA 423
Qy	109 -----TyrAlaProGluGlyValGlnLys 116
Db	424 GCTTTGCTCTCCACAGCCCTGGATGAGACCTCTGTTAGCACCAGCAAGGCGTCCAGAAG 483
Qy	117 IleLeuIleGlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGly 136
Db	484 ATCTTTATTTGGGAATAAGCTGTATGAGGAGCAGAAACGCGAGGTGGGAAGAGACAAGG 543
Qy	137 GlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeu 156
Db	544 CAGCAGCTGGCGAAGGAGTATGGATGACTTCTATGAACAAGTGCCTGCACCACTC 603
Qy	157 AsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu 176
Db	604 AACATTAAAGATCATTCACCGCTGTACAGAGCTGGTGTGTCAGGCCCATAGGAGAG 663
Qy	177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db	664 CTGGAAGGCTTCGGATGCGTGCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGAG 723
Qy	197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTyrCys 212
Db	724 GAGGGCAAAACCGAGGGCCCGAGCACTCTTCGAAACCTCTCGTGTGC 771
RESULT 5	
LOCUS	BC063736 2348 bp mRNA linear VRT 09-DEC-2003
DEFINITION	Xenopus laevis cDNA clone MGC:68722 IMAGE:4057038, complete cds.
ACCESSION	BC063736
VERSION	BC063736.1 GI:39645074
KEYWORDS	MGC.
SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
	Xenopodinae; Xenopus.
REFERENCE	1 (bases 1 to 2348)
AUTHORS	Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL	
PUBMED	12454917
2 (bases 1 to 2348)	
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenman,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Sapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Arancini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Wollaton,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmezel,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
12477932	
3 (bases 1 to 2348)	
Klein,S. and Strausberg,R.	
Direct Submission	
Submitted (08-DEC-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA	
NIH-MGC Project	
Contact: XGC help desk	
Email: cgapbs-xemail.nih.gov	
Tissue Procurement: Dr. Igor Dawid	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NIHC), Gaithersburg, Maryland	
Web site: http://www.nisc.nih.gov/	
Contact: nisc_mgc@nhgri.nih.gov	
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Series: IRAC Plate: 128 Row: 1 Column: 12	
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.	
Location/Qualifiers	
1. 2348	
/organism="Xenopus laevis"	
/mol_type="mRNA"	
/db_xref="taxon:8355"	
/clone="MGC:68722 IMAGE:4057038"	
/tissue_type="Lung, adult Xenopus"	
/clone_lib="NICHD_XGC_Lul"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
138. .569	
/codon_start=1	
/product="Unknown (protein for MGC:68722)"	
/protein_id="AAH63736.1"	
/db_xref="GI:39645075"	
/translation="MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNFPHPSHISTIGVD FKMTIEVDGIKVRIQIWDTAGQRYQTITKQYVYRAAGIPLVYDITRSYQIMKW	



```
/db_xref="GI:15489394"
/db_xref="LocusID:104886"
translation="MAKQYDVLFRLLILGDSGVGKTCLLCRFTDNFHFHSHISTIGVD
FKMTIDVDGKIKVRIQWDTAGQRYOTITKQYVRRAGQIFLVYDISSERSYQHMKW
VSDYDEVAPEGVQKILIGNADEQKRVQREGQQLAKEYGMDFYETSACFFSSPFS
PSTSPPL"
misc_feature
175..609
/note="RAB; Region: Rab subfamily of small GTPases"
/db_xref="CDD:cd00154"

ORIGIN
Alignment Scores:
Pred. No.: 7,38e-69 Length: 2560
Score: 798.00 Matches: 152
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.35% Mismatches: 0
Query Match: 72.22% Indels: 0
DB: 10 Gaps: 0

US-09-817-198c-2 (1-212) x BC013790 (1-2560)
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db 154 ATGCGGAACAGTAGATGCTGCTTCGGCTACTGCTGATCGGGGACTCCGGGGTTGGC 213
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 214 AAGACATGCTGCTGCTGCGCTTCACCGAACACGAGTTCCTCCTCGCATATCTCCACC 273
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 274 ATCGGTGTTGACTTTAAGATGAAGCATATCATGATAGACGGCATCAAAAGTGAGAATACAG 333
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 334 ATTGGGACACAGACGGCAGGAGGATGACAGACTATCAAAAGCAGTACTATCGGGCA 393
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 394 GCCCAGGGAATATTTTAGTCTAGCATATTAGCATGAGCGCTCTCTATCAGCATATCATG 453
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 454 AAGTGGGTGTCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 513
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValAlaGlyArgGluGlnGlyGlnGlnLeuAla 140
Db 514 AATAAGGCTGATGAAGAGCAAAACCGCAGGTGGGGGAGAGCAGCGGCAGCAGCTGGCT 573
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCys 153
Db 574 AAGAGTACGCGCATGGACTTCTACGAACAAGTGCCTGC 612

RESULT 7
HSM806937 2210 bp mRNA linear PRI 28-AUG-2003
LOCUS HSM806937
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J06205 (from clone DKFZp686J06205).
ACCESSION BX640825
VERSION BX640825.1 GI:34365090
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2210)
AUTHORS Bloecher,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
CONSRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
```

sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp686J06205) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES  
Location/Qualifiers  
1..2210  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="14q23.3"  
/clone="DKFZp686J06205"  
/tissue\_type="human fetal brain"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"  
/dev\_stage="fetal"  
1..2210  
/gene="DKFZp686J06205"  
<1..423  
/gene="DKFZp686J06205"  
/note="similarity to GTP-binding protein rab15 (Rattus norvegicus)"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAE45901.1"  
/db\_xref="GI:34365091"  
/translation="GPRPQGRNCQPLRLBAPEAGVRDKSRGVDFPMKTIIEVDGIK  
VRQIWDTAGQRYOTITKQYVRRAGQIFLVYDISSERSYQHMKWSDVDEVDGATS  
LPGCGGASPGKARRGPDGKANASRKLCLPQPMWTKTSG"

ORIGIN  
Alignment Scores:  
Pred. No.: 6,9e-69 Length: 2210  
Score: 797.50 Matches: 171  
Percent Similarity: 50.89% Conservative: 0  
Best Local Similarity: 50.89% Mismatches: 0  
Query Match: 72.17% Indels: 165  
DB: 9 Gaps: 1

US-09-817-198c-2 (1-212) x HSM806937 (1-2210)  
Qy 42 GlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIle 61  
Db 88 GGTGTTGACTTTAGATGAAGACCATAGAGTAGAGCGCATCAAAAGTCGGATACAGATC 147  
Qy 62 TrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgala 81  
Db 148 TGGGACACTGCAGGCGAGGAGATACAGACCATCAAAAGCAGTACTATCGCGGGCC 207  
Qy 82 GlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLys 101  
Db 208 CAGGGGATATTTTGGTCTATGATATAGCAGCGCGCTCTTACCAGCATCATCAAG 267  
Qy 102 TrpValSerAspValAsp----- 107  
Db 268 TGGGTGAGTGCATGATGAGGTAGGATGAGTCCACCTCATCTGCGGGGTGTGGAGAGGT 327  
Qy 107 ----- 107  
Db 328 GCCTCAGCGGGGAGGCAAGCGGAGGCCAGATGGGAAGCAANTGCTCCAGGAGCTT 387  
Qy 107 ----- 107  
Db 388 TGCCTTCCACAGCCCTGGATGAAGACCTCTGGGTGAGTAAGACATCGGGAAGAAACCGAA 447  
Qy 107 ----- 107  
Db 448 GCTGCCATGCCCTCACTCTCTATACCTGCGAGGCGCTCCACGGGTGTGTCTTTCCCGGAA 507  
Qy 107 ----- 107



```
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 575 GTCTGGGACACAGCAGGACAGGAGGTTTAAGACCATCACCAGTGCATCTACAGAGG 634
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 635 GCCATGGGAATCATCTCGTGTATGACATAATCTATGAAAATCTATGAGAAATTCAG 694
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 695 AACTGGATGAAGACATCAAGAAATGATCAGCGGTGTAACTGCAATGTTACTGGT 754
Qy 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnLysAla 140
Db 755 AACAAAGTGTGATTCAGGCGCAAGAGAAAGTCTCAAAAGAGAGCTGGTGAAGCTCGCA 814
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 815 AAGAAACATGTTATTCGATTCGAGACATTTGCTGAAGTCAAAATAGAGCCGGTTCAT 874
Qy 161 SerPheThrArgLeuThr---GluLeuValLeuGlnAlaHisArgLys-----Glu 176
Db 875 TCTTTTACTCTCTGTCGAGACATTTGCTGAAGTCAAAATAGAGCCGGTTCATCT 934
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db 935 GGACGTGAGGTCAAACTTACCAGCAGACA-GAAAAAGTCTCTCTCCCAAAATGCTCTTCT 993
Qy 197 GluGlyLysPro 200
Db 994 TTAGGCAAAACCA 1005

RESULT 9
AB006189 1613 bp mRNA linear INV 07-AUG-1997
LOCUS Drosophila melanogaster mRNA for Rab10, complete cds.
ACCESSION AB006189
VERSION AB006189.1 GI:2317271
KEYWORDS Rab10.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (sites)
AUTHORS Satoh,A.K., Tokunaga,F. and Ozaki,K.
TITLE Rab proteins of Drosophila melanogaster: novel members of the
Rab-protein family
JOURNAL FEBS Lett. 404 (1), 65-69 (1997)
MEDLINE 97228579
PUBMED 9074639
REFERENCE 2 (bases 1 to 1613)
AUTHORS Ozaki,K.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1997) Koichi Ozaki, Osaka University, Graduate
School of Science, Department of Biology; 1-1 Machikaneyama,
Toyonaka, Osaka 560, Japan (E-mail:ozaki@bio.sci.osaka-u.ac.jp,
Tel:+81-6-850-5439, Fax:+81-6-850-5439)
FEATURES
source
1. 1613
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Oregon R"
/db_xref="taxon:7227"
/tissue_type="head"
1. 1613
/gene="DRab10"
257..871
/gene="DRab10"
/codon_start=1
/product="Rab10"
/protein_id="BAA21744.1"
```

```
/db_xref="GI:2317272"
/translation="MAKTYDYLKLLKLLIGDSGVGKTCILFRFSDDAFTSTFTIGI
DFKIKTVLRGKKIKLQIDWTAGQERFHTTTSYRGAMGLMVDYITNEKSPFNIVK
WLRNIDSHANDEKEMILGNKCDMTDKRVNKRGEALAREHGRFMTSAKSNINIE
RAFCELAELIDKTSRESAENQERVIIDRRNQEKAPGYSKCCA"

ORIGIN
Alignment Scores:
Pred. No.: 2,52e-44 Length: 1613
Score: 546.50 Matches: 105
Percent Similarity: 72.04% Conservative: 47
Best Local Similarity: 49.76% Mismatches: 50
Query Match: 49.46% Indels: 9
DB: 3 Gaps: 3

US-09-817-198C-2 (1-212) x AB006189 (1-1613)

Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
Db 257 ATGCCAAGAAACCTACGATTTGCTCTTTAAACTGTGCTGATTCGGTGATTCAGGAGTG 316
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 317 GGCAGACAGTCATATTGTTCCGGTTCTCGATGATGCATTCAGTCCACGTTCAATTCG 376
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
Db 377 ACCATAGCATCGATTTCAAATCAAAACAGTCGAGTCGCGCGCAAGAAGATCAAGCTG 436
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
Db 437 CAAATATGGGACACCGCGCGCAGGCGGTTCACACAGATAACCAACCCCTCTACTATCGA 496
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 497 GCGCCCATGGGCATATGCTGCTATGACATAACGAACAGAGAGAGTTTCGAGAACATA 556
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 557 GTCAATGGTTACGGAATATAGACAGACACGCCAACGAGGATGTGGAGAAGATGATCCTC 616
Qy 120 GlyAsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnLys 139
Db 617 GGCACAAGTCCGATATACGCGACAGAGGGTGTCAACAGGAGCGCGCGCAAGCGATT 676
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 677 GCGCGTGAACATGGCATTCGGTTTATGAAACATCCGCAAGTCCGAACATAAATCAG 736
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGly 179
Db 737 CGGCGCTTCTCGAGCTGGCGAGGCCATCTGGAC-----AAGACATCAGGA 784
Qy 180 LeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLys 199
Db 785 CGCGAGTCGCGGAGAATCAGAGCGGTGATTTATTCATCGCCGCAACAGGAGGAGCG 844
Qy 200 ProGluGlyProAlaAsnSerSerLysThrCys 210
Db 845 CCG-----GGCTACAGCAAGTGCTGC 865

RESULT 10
AB006425
LOCUS Drosophila melanogaster LD39986 full length cDNA.
DEFINITION Drosophila melanogaster LD39986 full length cDNA.
ACCESSION AB006425
VERSION AB006425.1 GI:16648397
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
```









AUTHORS Zahraoui,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-NOV-1990) A. Zahraoui, INSERM-U 248, 10 AVENUE DE VERDUN, 750-10 PARIS, FRANCE

FEATURES  
 source Location/Qualifiers  
 1..660 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Human pheochromocytoma cDNA library"  
 gene 1..660 /gene="rab8"  
 CDS 11..634 /gene="rab8"  
 /codon\_start=1  
 /product="rab8 small GTP binding protein"  
 /protein\_id="CAA40065.1"  
 /db\_xref="GI:452318"  
 /db\_xref="GOA:P24407"  
 /translation="MAKTYDYLPLKLLIGDSGVGKTCVLPFRFSDAFNSTFISTIGID FKRTIELDGKRIKLIQWDTAGOERPTITAYVRGAMGLMVDITNEKSFDIRNW IRNIEERASADVEKMLGNKCDVNDKROVSKERGEKALDYGIKFMETSAKANINVEN AFPTLARDIKAKMDKLEGNSPQSGNQGKVTTPDQKRSSFRCVLL"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1..186-44 Length: 660  
 Score: 545.00 Matches: 99  
 Percent Similarity: 77.01% Conservative: 45  
 Best Local Similarity: 52.94% Mismatches: 43  
 Query Match: 49.32% Indels: 0  
 DB: 9 Gaps: 0

US-09-817-198C-2 (1-212) x HSMRAB8 (1-660)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuLeuLeuLeuValGly 20  
 Db 11 ATGGCGAAGACCTACGATTACCTGTTCAAGCTGCTGCTGATCGGGGCTCGGGGTGGG 70  
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40  
 Db 71 AAGACCTGTGTCTTCCTGCTTCGAGAGACGCTTCAACTCCACTTTATCTCCACC 130  
 Qy 41 IleGlyValAspPheLeuMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
 Db 131 ATAGGAATTGACTTTAAATTTAGGACCATAGACTCGATGCAAGAGATTAATTAATCGAG 190  
 Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
 Db 191 ATATGGGACACAGCGGTGAGGACGGTTTCGAGAGATCAACACGGCTACTACAGGGGT 250  
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
 Db 251 GCAATGGGCATCATCTGCTTACGACATCAACACGAGAGAGTCTCTCGACACATCCGG 310  
 Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
 Db 311 AACTGGGATTCGCAATGAGGAGCAGCGCTCTCGAGACGTCGAAAGATGATATATCGGG 370  
 Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLeuAla 140  
 Db 371 AACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 430  
 Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 Db 431 CTCGACTATGGAATCAAGTTCATGAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 490  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 Db 491 GCATTTTCTCTCGCCAGAGATATCAAGAGCAAAATGGACAAAAATTCGAGGCAAC 550  
 Qy 181 ArgMetArgAlaSerAsnGlu 187

Db 551 AGCCCCAGGGGAGCACACAG 571  
 RESULT 15  
 CFRAB8  
 LOCUS Canine rab8 mRNA for ras-related GTP-binding protein. 760 bp mRNA linear MAM 30-SEP-1999  
 DEFINITION X56385  
 ACCESSION X56385  
 VERSION GI:920  
 KEYWORDS GTP-binding protein; protein transport; ras gene; ras gene family.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 Chavrier,P., Vingron,M., Sander,C., Simons,K. and Zerlial,M.  
 Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line  
 JOURNAL Mol. Cell. Biol. 10 (12), 6578-6585 (1990)  
 MEDLINE 91061765  
 PUBMED 2123294  
 REFERENCE 2 (bases 1 to 760)  
 AUTHORS Zerlial,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-1990) Zerlial M., EMBL, Meyerhofstrasse 1, 6900 Heidelberg, FRG

FEATURES  
 source Location/Qualifiers  
 1..760 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /strain="cocker spaniel"  
 /db\_xref="taxon:9615"  
 /clone="2"  
 /cell\_line="MDCK"  
 gene 1..760 /gene="rab8"  
 CDS 10..633 /gene="rab8"  
 /functions="GTP binding protein"  
 /codon\_start=1  
 /protein\_id="CAB56776.1"  
 /db\_xref="GI:6006436"  
 /db\_xref="GOA:P24407"  
 /db\_xref="SWISS-PROT:P24407"  
 /translation="MAKTYDYLPLKLLIGDSGVGKTCVLPFRFSDAFNSTFISTIGID FKRTIELDGKRIKLIQWDTAGOERPTITAYVRGAMGLMVDITNEKSFDIRNW IRNIEERASADVEKMLGNKCDVNDKROVSKERGEKALDYGIKFMETSAKANINVEN AFPTLARDIKAKMDKLEGNSPQSGNQGKVTTPDQKRSSFRCVLL"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1..41e-44 Length: 760  
 Score: 545.00 Matches: 99  
 Percent Similarity: 77.01% Conservative: 45  
 Best Local Similarity: 52.94% Mismatches: 43  
 Query Match: 49.32% Indels: 0  
 DB: 4 Gaps: 0

US-09-817-198C-2 (1-212) x CFRAB8 (1-760)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuLeuLeuValGly 20  
 Db 10 ATGGCGAAGACCTACGATTACCTGTTCAAGCTGCTGCTGATCGGGGCTCGGGGTGGG 69  
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40  
 Db 70 AAGACCTGTGTCTTCCTGCTTCGAGAGACGCTTCAACTCCACTTTATCTCCACT 129  
 Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
 Db 130 ATAGGAATTGACTTTAAATTTAGGACCATAGACTCGATGCAAGAGATTAATTAATCGAG 189  
 Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80

```
Db 190 ATATGGGACACAGCTGTCTCAAGAACGGTTTCGGGACGATCACACAGCCTATTACAGGGGC 249
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 250 GCAATGGGCATCATGCTGGTCTATGACATCACCAACGAGAAATCCTTTGACAAATATCCGG 309
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 310 AACTGGGATTCGGAACATGAGGACATGCTTCTCCAGATGTCGAAAGATGATACTCGGA 369
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140
Db 370 AACAAAGTGTGATGTAACGACAAAGACAAAGTTTCCAAAGGAACGGGAGAAAGCTGCC 429
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 430 CTGGACTATGGAATCAAGTTTCATGGAGACCAGTCCGAAGGCCAACATCAATGTGGAGAAC 489
Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 490 GCATTTTCTCTTCGCCAGAGACATCAAGCAAGATGGACAAAAAATTGGAAGGCCAAC 549
Qy 181 ArgMetArgAlaSerAsnGlu 187
Db 550 AGTCCCCAAGGGAGCAACCAG 570
```

Search completed: March 21, 2004, 12:57:23  
Job time : 2753 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:01:53 ; Search time 354 Seconds  
(without alignments)  
2544.119 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAQYDVLFRLLIGDSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09817198/runat\_15032004\_101745\_20514/app\_query.fasta\_1.391  
-DB=N Geneseq\_29Jan04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN\_1\_1\_708 @runat\_15032004\_101745\_20514 -NCFU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- N\_Geneseq\_29Jan04.\*
- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	1054	6 ABQ93361	Abq93361 Human cDN
2	1105	100.0	2021	4 AAS27053	Aas27053 cDNA enco
3	1105	100.0	2021	9 ADB93231	Adb93231 Human cDN
4	1105	100.0	3257	7 AAD47168	Aad47168 Human Ras
5	1092	98.8	1666	3 AAC75813	Aac75813 Human ORF
6	832	75.3	566	4 AAS27472	Aas27472 cDNA enco
7	832	75.3	566	9 ADB93650	Adb93650 Human cDN
8	546.5	49.5	1540	4 ABL27707	AbL27707 Drosophil

9	545	49.3	1986	4	ABA09160	Aba09160 Human rab
10	540	48.9	2411	6	ABQ54410	Abq54410 Human ova
11	534.5	48.4	1274	3	AAFI18072	Aafi18072 Lung canc
12	530.5	48.0	716	3	AAA40104	Aaa40104 Human Rab
13	530.5	48.0	861	3	AAA40108	Aaa40108 Human Rab
14	530.5	48.0	888	3	AAA40109	Aaa40109 Canine Ra
15	530.5	48.0	956	3	AAA96887	Aaa96887 Nucleotid
16	530.5	48.0	1537	3	AAFI16196	Aafi16196 Human pro
17	530.5	48.0	3164	6	ABX04179	Abx04179 Human mRN
18	530.5	48.0	3533	4	AAH17889	Aah17889 Human CDN
19	529.5	47.9	674	5	AAS71453	Aas71453 DNA encod
20	528.5	47.8	1161	4	AAH75182	Aah75182 Nucleotid
21	528.5	47.8	1265	6	ABV78060	Abv78060 Hypoxia-r
22	528.5	47.8	2247	4	ABL29661	AbL29661 Drosophil
23	528.5	47.8	2497	4	AAS60878	Aas60878 Human can
24	528.5	47.8	2497	4	AAS60884	Aas60884 Human can
25	528.5	47.8	2497	4	AAS60893	Aas60893 Human can
26	528.5	47.8	2497	4	AAS60895	Aas60895 Human can
27	528.5	47.8	2497	5	ABV25781	Abv25781 Human pro
28	528.5	47.8	2497	5	ABV30037	Abv30037 Human pro
29	528.5	47.8	3077	4	AAH13912	Aah13912 Human CDN
30	523.5	47.4	775	9	ADC76178	Adc76178 DNA homol
31	523.5	47.4	775	9	ADD17377	AdD17377 DNA (SeqI
32	523	47.3	621	7	ABT21250	Abt21250 Aspergill
33	523	47.3	621	7	ABT19071	Abt19071 Aspergill
34	519.5	47.0	866	4	AAH04301	Aah04301 Human cDN
35	519	47.0	651	6	AB213031	Ab213031 Arabidops
36	519	47.0	1023	3	AAC51491	Aac51491 Arabidops
37	519	47.0	1023	3	AAC34347	Aac34347 Arabidops
38	513.5	46.5	1129	3	AAC44482	Aac44482 Zea may
39	510.5	46.2	572	9	ADC76238	Adc76238 DNA homol
40	510.5	46.2	572	9	ADD17370	AdD17370 DNA (SeqI
41	510	46.2	911	3	AAC38429	Aac38429 Arabidops
42	509.5	46.1	709	9	ADD16498	AdD16498 DNA (SeqI
43	509	46.1	988	9	ADC75639	AdC75639 DNA homol
44	509	46.1	988	9	ADD16501	AdD16501 DNA (SeqI
45	507.5	45.9	720	9	ADD16526	AdD16526 DNA (SeqI

ALIGNMENTS

RESULT 1  
ABQ93361  
ID ABQ93361 standard; cDNA; 1054 BP.

AC ABQ93361;

DT 14-OCT-2002 (first entry)

DE Human cDNA SEQ ID NO 74.

XX Human; vulnary; dermatological; neuroprotective; nootropic; cancer;  
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;  
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
KW burn; central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; immune disorder;  
KW autoimmune disorder; multiple sclerosis; diabetes; allergy; gene; ss.

OS Homo sapiens.

XX WO200218424-A2.

XX 07-MAR-2002.

XX 31-AUG-2001; 2001WO-US027093.

XX 01-SEP-2000; 2000US-00654935.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;  
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;



PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241825P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI: 2001-465460/50.  
DR P-PSDB; AAU17136.  
DR  
XX  
PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders.  
XX  
XX Claim 1; SEQ ID NO 88; 880pp; English.  
PS  
XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative disorders  
CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
CC respiratory disorders, dermatological disorders, in wound healing,  
CC epithelial cell proliferation, endocrine disorders (e.g. Addison's  
CC disease), reproductive system disorders, gastrointestinal disorder  
CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
CC B-cell responsiveness to pathogens, activators of T-cells, to induce  
CC higher affinity antibodies, and as a means to induce tumour proliferation  
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-  
CC AAS27850 represent novel signal transduction pathway protein coding  
CC sequences and PCR primers of the invention  
XX

Alignment Scores:  
Pred. No.: 8.95e-118 Length: 2021  
Score: 1105.00 Matches: 212  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0















Db 251 ATCTGGGACACTGCAGGGCAGGAGATACACAGACCATCAACAGCAGTACTATCGCGG 310  
Qy 81 AlaGlnGlyLeuPheLeuValTyArgPheLeuSerSerGluArgSerTyArgGlnHisLeuMet 100  
Db 311 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGCGCTCTTACCAGCACATCATG 370  
Qy 101 LysTrpValSerAspValAspGluTyArgAlaProGluGlyValGlnHisLeuLeuGly 120  
Db 371 AAGTGGGTCTAGTACGTGGATGAGTACGCACACAGAGCGCTCCAGAAATCTTATTGGG 430  
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnLysLeuAla 140  
Db 431 AATAAGGCTGATCAGGAGCAGACGAGACGCGAGTGGGAAGAGAGCAAGGCGCAGCAGCTGGCG 490  
Qy 141 LysGluTyArgMetAspPheTyArgGlnThrSerAlaCysThrAsnLeuAsnLeuGly 160  
Db 491 AANAGAGTATGGCATGGGACTTCTATGAAACAAGTGCTGCACCACTTAAACATTAAGAG 550  
Qy 161 SerPheThrArg 164  
Db 551 TCATTCACGCGT 562

RESULT 8

ABL27707  
ID ABL27707 standard; DNA; 1540 BP.  
AC ABL27707;  
XX  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 34594.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Claim 1; SEQ ID NO 34594; 21pp + Sequence Listing; English.

XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO.int/pub/published\_pct\_sequences

XX  
SQ Sequence 1540 BP; 437 A; 353 C; 420 G; 330 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.3e-53 Length: 1540

Score: 546.50 Matches: 105  
Percent Similarity: 72.04% Conservative: 47  
Best Local Similarity: 49.76% Mismatches: 50  
Query Match: 49.46% Indels: 9  
DB: 4 Gaps: 3  
US-09-817-198c-2 (1-212) x ABL27707 (1-1540)  
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuGlyAspSerGlyVal 19  
Db 257 ATGGCAAGAAACACCTACGATTTGCTCTTTAACTGTTCTGATCGTTCATTCAGGAGTG 316  
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSer 39  
Db 317 GGCACAGCTGCATATTGTTCCGGTTCTCGGATGATGCATTCACGTCCACGTTTCATATCG 376  
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysLeuValArgIle 59  
Db 377 ACCATAGGCATCGATTTCAAAATCAAAACAGTCGAGCTGCGCGCAAGAGATCAAGCTG 436  
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyArgGlnThrIleThrLysGlnTyArg 79  
Db 437 CAATATGGACACCCCGCGCAGGAGCGGTTCCACAGATAACCACTCGTACTATCGA 496  
Qy 80 ArgAlaGlnGlyLeuPheLeuValTyArgPheLeuSerSerGluArgSerTyArgHisLeu 99  
Db 497 GCGCGCATGGCATAATGCTGCTATGACATCAATCAACGAGAGAGAGTTTCGAGAACATA 556  
Qy 100 MetLysTrpValSerAspValAspGluTyArgProGluGlyValGlnLysLeuLeu 119  
Db 557 GTCAAATGGTTACGGAATATAGACGAGCAGCGCAACGAGATGTGGAGAGATGATCCTC 616  
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139  
Db 617 GGCACACAGTGCATATGACGACAGAGGGTGTCAACAGGAGCGCGCGAAGCGATT 676  
Qy 140 AlaLysGluTyArgMetAspPheTyArgGluThrSerAlaCysThrAsnLeuAsnLeuLys 159  
Db 677 GCGCGTGAACATGGCATTCGGTTTATGGAACATCCGCCAAGTCGAACTAATACATCGAG 736  
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGly 179  
Db 737 CGGCGCTTCTGCGAGCTGCGCGGAGCGCATTTCTGGAC-----AAGACATCAGGA 784  
Qy 180 LeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLys 199  
Db 785 CGCGAGTCGCGGAGAAATCAGGAGCGGTGATTATCGATCGCCGGAACAGGAGAGCGG 844  
Qy 200 ProGluGlyProAlaAsnSerSerLysThrCys 210  
Db 845 CCG-----GGCTACAGCAAGTGCTGC 865

RESULT 9

ABA09160  
ID ABA09160 standard; cDNA; 1986 BP.  
XX  
XX ABA09160;  
XX  
XX 11-JAN-2002 (first entry)  
XX  
XX Human rab8 homologue-encoding cDNA, SEQ ID NO:936.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 XX antifungal; vulnerary; antiulcer; ss.

OS Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003800.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX P-PSDB; ABB11916.

XX Human proteins and DNA encoding sequences useful for preventing, treating  
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis  
 PT and cancer.

XX Claim 1; Page 805; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention

SQ Sequence 1986 BP; 519 A; 545 C; 469 G; 453 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.95e-53	Length:	1986
Score:	545.00	Matches:	99
Percent Similarity:	77.01%	Conservative:	45
Best Local Similarity:	52.94%	Mismatches:	43
Query Match:	49.32%	Indels:	0

DB: 4 Gaps: 0  
 US-09-817-198C-2 (1-212) x ABA09160 (1-1986)

Qy 1 MetalalysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20  
 Db 19 ATGGCGAAGACCTACGATTACCTGTTCAAGCTGCTGATCGGGGACTCGGGGGTGGG 78  
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40  
 Db 79 AAGACCTGTGCTGCTGCTTCGCGAGGACGCTTCACTCCACTTTTATCTCCACC 138  
 Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
 Db 139 ATAGGAATTGACTTTAAATAATTAGGACCATAGAGCTCGATGGCAAGAGAAATTAACACTGCAG 198  
 Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArg 80  
 Db 199 ATATGGGACACAGCCGCTCAGGAACGGTTTCGGACGATCAACCGGCTACTACAGGGGT 258  
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
 Db 259 GCATGGGACATCATGCTGGTCTAGCATATCACCACGAGAGAGTCTTCGACACATCCGG 318  
 Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
 Db 319 AACTGGATTGCAACATTTAGGAGCAGCCTCTCGACAGCTCGAAGAAAGATGATACTCGG 378  
 Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGlnGlnGlyGlnGlnLeuAla 140  
 Db 379 AACAAAGTGTATGTAATGACAGAGACAAGTTTCCAAAGGAACGGGAGAGAAAGCTGGCC 438  
 Qy 141 LysGlyTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 Db 439 CTCGACTATGGAATCAAGTTTCATCGACACAGCGCGAGGCCACATCAATGTGAAAAAT 498  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 Db 499 GCATTTTTCATCTCTCGCCAGAGATATCAAAAGCAAAAATGGACAAAAAATTTGGAAGCAAC 558  
 Qy 181 ArgMetArgAlaSerAsnGlu 187  
 Db 559 AGCCCCCAGGGGAGCAACCCAG 579

RESULT 10

ABQ54410  
 ID ABQ54410 standard; cDNA; 2411 BP.

XX AC ABQ54410;

XX 22-AUG-2002 (first entry)

DE Human ovarian antigen HCGMA67 cDNA, SEQ ID NO:290.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; gene; ss.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX

```

PR 07-JUN-2000; 2000US-0209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
DR P-PSDB; ABP41333.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
XX Claim 1; SEQ ID NO 290; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2411 BP; 600 A; 671 C; 583 G; 547 T; 0 U; 10 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.44e-52 Length: 2411
XX Score: 540.00 Matches: 98
XX Percent Similarity: 76.88% Conservative: 45
XX Best Local Similarity: 52.69% Mismatches: 43
XX Query Match: 48.87% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-817-198C-2 (1-212) x ABQ54410 (1-2411)
XX
XX 2 AlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGlyLys 21
XX
XX 47 GCGAAGACCTACGATTACCTCTTCAAGCTGCTGATCGGGGACTCGGGGGTGGGAAG 106
XX
XX 22 ThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSerThrIle 41
XX
XX 107 ACCTGTGTCTTGTTCGCGCTTCTCCGAGACGCCCTTCACTCCATTTATCTCCACCATA 166
XX
XX 42 GlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGlnIle 61
XX
XX 167 GGAATTGACTTTAAATTAAGACCATAGAGCTGATGCGAAGAGTAATTAACCTCAGATA 226
XX
XX 62 TrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArgAla 81
XX
XX 227 TGGGACACAGCCGCTCAGAACCGGTTTCGGACGATCACAACGGGCTACTACAGGGGTGCA 286
XX

```

```

QY 82 GlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLys 101
XX
XX 287 ATGGGCATCATGCTGCTTACGACATCACCAACGAGAAGTCTCTCGACAACATCCGGAAC 346
XX
XX 102 TrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsn 121
XX
XX 347 TGGATTTCGCAACATTGAGGAGCAGCGCTCTGCAGACGTCGAAAAGATGATATCTCGGGAAC 406
XX
XX 122 LysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLys 141
XX
XX 407 AAGTGTGATGTGAATGACAGAGACAAGTTTCCAGGAACGGGGAGAAAGCTTGGCCCTC 466
XX
XX 142 GluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSer 161
XX
XX 467 GACTATGGGAATCAAGTTTCATGGAGACCGCGGAGCCCAACATCAATGTGGAATAATGCA 526
XX
XX 162 PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGlyLeuArg 181
XX
XX 527 TTTTTCACCTCTCGCCAGAGATATCAAAGCAAAAATGGACAAAATAATTGGAAGGCAACAGC 586
XX
XX 182 MetArgAlaSerAsnGlu 187
XX
XX 587 CCCAGGGGAGCAACACG 604
XX
XX RESULT 11
XX AAF18072
XX ID AAF18072 standard; DNA; 1274 BP.
XX AC AAF18072;
XX XX
XX DT 14-MAR-2001 (first entry)
XX XX
XX DE Lung cancer associated polynucleotide sequence SEQ ID 91.
XX
XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;
XX cardioactive; immunomodulatory; muscular active; vulnerary;
XX gastrointestinal; nephrotropic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.
XX
XX OS Homo sapiens.
XX
XX PN WC2000055180-A2.
XX
XX PD 21-SEP-2000.
XX
XX PF 08-MAR-2000; 2000WO-US005918.
XX
XX PR 12-MAR-1999; 99US-0124270P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX
XX PI Ruben SM;
XX
XX DR WPI; 2000-587514/55.
XX DR P-PSDB; AAB58196.
XX
XX XX Lung cancer associated gene sequences, referred to as lung cancer
XX PT antigens, useful for treatment, prevention, and diagnosis of disorders
XX PT such as lung cancer.
XX
XX PS Claim 1; Page 566; 1425pp; English.
XX
XX CC Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer
XX CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX CC associated proteins and polynucleotide sequences, their agonists, and
XX CC antagonists may have neuroprotective; cytostatic; cardioactive;
XX CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
XX CC general; nephrotropic; antiinfective; gynecological; or antibacterial
XX CC activity. The invention also includes antibodies specific for the protein
XX CC or polynucleotide sequences. The lung cancer associated polynucleotide

```







```

DB: 3 Gaps: 1
US-09-817-198c-2 (1-212) x AAA40109 (1-888)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 133 ATGCGAAGAACGAGTACGACCTCTTTCAAGCTGCTCTCGATCGGAGACTCGGGAGTA 192
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 193 GGGNAGACCTGGCGCTCTTTTCGTTTCGGATGATCCCTTCATACACCACTTTATTTC 252
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIle 59
Db 253 ACCATAGGAATAGATTTTAAGATCAAAACAGTTGAATTACAAGGAAAGAGATCAAGCTA 312
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
Db 313 CAGATATGGGATACAGAGCGCCAGGCGGATTTTCCACCAATCAACCACTCTCTACTAGA 372
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 373 GGACATGGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAAGACATC 432
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeu 119
Db 433 AGCAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAGAAGATGTTACTA 492
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeu 139
Db 493 GGAACAAATGTGATATGGACGATAAAGAGTTGTCTACCTAAAGGAAAGGAGACGAT 552
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 553 GCAAGGAGCATGTTAGTATTTTGGAGCTAGTGCAGAAAGTAAATATAAATCAATCGAA 612
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 613 AAGGCTTCTCTCACATTAGCTGAAGATATCTCTCGA 648

RESULT 15
AAA96887
ID AAA96887 standard; DNA; 956 BP.
XX
AC AAA96887;
XX
DT 19-FEB-2001 (first entry)
XX
DE Nucleotide sequence of human RAB10.
XX
KW RAB protein; GTPase; GTP binding; gene therapy; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 127..729
FT /*tag= a
FT /product= "RAB10"
XX
XX WO200058464-A2.
XX
XX 05-OCT-2000.
XX
XX 13-MAR-2000; 2000WO-US006330.
XX
XX 25-MAR-1999; 99US-0126083P.
XX
XX (AXYS-) AXYS PHARM INC.
XX
XX Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
XX WPI; 2000-647233/62.
XX P-PSDB; AAB19165.

```

```

XX
PT Novel isolated nucleic acid encoding a mammalian RAB protein useful for
PT identifying homologous or related genes, in producing composition that
PT modulates expression or function of RAB for cancer therapy.
XX
PS Claim 4; Page 38-39; 58pp; English.
XX
CC The present sequence encodes a mammalian RAB protein. RAB proteins
CC constitute the largest family of small GTPases, with over 40 identified
CC isoforms. RAB proteins contain four highly conserved peptide sequences
CC involved in GTP binding and hydrolysis. Compositions comprising RAB
CC nucleic acid are useful for identifying homologous or related genes, in
CC producing compositions that modulate the expression or function of RAB,
CC for gene therapy, mapping functional regions of the protein and in
CC studying associated physiological pathways. In addition, modulation of
CC the gene activity in vivo is used for prophylactic and therapeutic
CC purposes, such as treatment of cancer, and identification of cell type
CC based on expression. The DNA may also be used to identify expression of
CC the gene in a biological specimen
XX
SQ Sequence 956 BP; 294 A; 205 C; 213 G; 244 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1.23e-51 Length: 956
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservative: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 3 Gaps: 1

US-09-817-198c-2 (1-212) x AAA96887 (1-956)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 127 ATGCGAAGAACGAGTACGACCTCTTTCAAGCTGCTCTCGATCGGAGACTCGGGAGTG 186
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 187 GGGNAGACCTGGCGCTCTTTTCGTTTCGGATGATCCCTTCATACCACTTTATTTC 246
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIle 59
Db 247 ACCATAGGAATAGATTTTAAGATCAAAACAGTTGAATTACAAGGAAAGAGATCAAGCTA 306
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
Db 307 CAGATATGGGATACAGAGCGCCAGGCGGATTTTCCACCAATCAACCACTCTCTACTAGA 366
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 367 GGGCAATGGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAAGACATC 426
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeu 119
Db 427 AGCAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAGAAGATGTTACTA 486
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeu 139
Db 487 GGAACAAATGTGATATGGACGATAAAGAGTTGTCTACCTAAAGGAAAGGAGACAGATT 546
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 547 GCAAGGAGCATGTTAGTATTTTGGAGCTAGTGCAGAAAGTAAATATAAATCAATCGAA 606
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 607 AAGGCTTCTCTCACATTAGCTGAAGATATCTCTCGA 642

```

Search completed: March 21, 2004, 12:11:24  
Job time : 362 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 04:59:32 ; Search time 83 Seconds  
(without alignments)  
1417.465 Million cell updates/sec

Title: US-09-817-198C-2  
Perfect score: 1105  
Sequence: 1 MAKQYDVLRLLLIGDSVG.....LEEEKGPEGANSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n\_model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US09817198/runat\_15032004\_101743\_20457/app\_query.fasta\_1.391  
-DB=issued Patents NA -QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN 1 1.85 @runat\_15032004\_101743\_20457 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478.5	43.3	925	2	US-08-916-901-4
2	478.5	43.3	925	4	US-09-154-602-4
3	477.5	43.2	730	4	US-09-300-958A-16
4	467	42.3	723	4	US-09-016-434-1422
5	467	42.3	8137	4	US-09-566-921-7
6	456.5	41.3	1069	4	US-09-620-312D-646
7	437.5	39.6	1340	2	US-08-824-873-2
8	437.5	39.6	1340	3	US-09-198-184-2
9	429	38.8	875	4	US-09-075-454-10
10	429	38.8	1106	4	US-09-620-312D-959
11	429	38.8	2612	4	US-09-484-970B-142
12	424	38.4	833	4	US-09-620-312D-426

13	424	38.4	842	4	US-09-255-920A-6
14	419	37.9	639	4	US-09-399-913-66
15	411.5	37.2	970	3	US-08-888-077A-28
16	404.5	36.6	3936	4	US-09-919-172-49
17	388.5	35.2	1054	4	US-09-976-594-1096
18	374	33.8	1102	4	US-09-620-312D-828
19	373	33.8	847	2	US-08-773-423-4
20	366	33.1	944	4	US-09-016-434-1060
21	343.5	31.1	3745	4	US-09-976-594-387
22	343	31.0	890	3	US-08-741-411-4
23	340.5	30.8	803	4	US-09-075-454-13
24	339	30.7	912	4	US-09-016-434-1159
25	332.5	30.1	1984	4	US-09-023-655-7
26	321.5	29.1	1407	4	US-09-493-914-1
27	320	29.0	1172	4	US-09-075-454-8
28	320	29.0	1533	4	US-09-075-454-11
29	320	29.0	1886	4	US-09-620-312D-647
30	313	28.3	1255	2	US-08-766-551-6
31	297	26.9	606	4	US-09-016-434-1082
32	297	26.9	1175	2	US-08-773-423-6
33	296	26.8	921	4	US-09-016-434-1124
34	286.5	25.9	464	4	US-09-300-958A-14
35	283.5	25.7	624	4	US-09-016-434-1083
36	282.5	25.6	1084	4	US-09-673-395A-138
37	282	25.5	848	3	US-08-741-411-2
38	278.5	25.2	1393	4	US-09-391-741A-3
39	277	25.1	1631	4	US-09-620-312D-587
40	276.5	25.0	1393	4	US-09-391-741A-17
41	274.5	24.8	1749	4	US-09-149-476-54
42	273.5	24.8	1393	4	US-09-391-741A-27
43	272	24.6	1127	4	US-09-391-741A-1
44	271.5	24.6	820	3	US-08-741-411-6
45	271.5	24.6	820	4	US-09-016-434-112

ALIGNMENTS

RESULT 1  
US-08-916-901-4  
; Sequence 4, Application US/08916901  
; Patent No. 5892012  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/916,901  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0367 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166

Sequence 6, Appli  
Sequence 66, Appl  
Sequence 28, Appl  
Sequence 49, Appl  
Sequence 1096, Ap  
Sequence 828, App  
Sequence 4, Appli  
Sequence 1060, Ap  
Sequence 387, App  
Sequence 4, Appli  
Sequence 13, Appl  
Sequence 1159, Ap  
Sequence 7, Appli  
Sequence 1, Appli  
Sequence 8, Appli  
Sequence 11, Appl  
Sequence 647, App  
Sequence 6, Appli  
Sequence 1082, Ap  
Sequence 6, Appli  
Sequence 1124, Ap  
Sequence 14, Appl  
Sequence 1083, Ap  
Sequence 138, App  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 587, App  
Sequence 17, Appl  
Sequence 54, Appl  
Sequence 27, Appl  
Sequence 1, Appli  
Sequence 6, Appli  
Sequence 112, App



Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140  
Db 426 AACAGAGCGACCTCACCACCAGAGGTGGTGACCAACACCAGCAAGGATTTGCA 485  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 486 GACTCTCTGGGCATCCCTCTTCTGGAGCAGCGCCCAAGATGCCCAATGTGCGACG 545  
Qy 161 SerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGluLeu 180  
Db 546 GCGTTC-----ATGACCATGGCTGCTGAATCAAAAAGCGGATGGGCGCT 590  
Qy 181 ArgMetArgAlaSerAsnGluLeuAlaGluLeuGluGluGluGlyLeuPro 200  
Db 591 GGAGCAGCTCTGGGGCGAGCGGCCCAATCTCAAGATCGACAGCACCCCTGTAAAGCG 650  
Qy 201 GluGly 202  
Db 651 GCTGGC 656  
RESULT 3  
US-09-300-958A-16  
; Sequence 16, Application US/09300958A  
; Patent No. 6495319  
; GENERAL INFORMATION:  
; APPLICANT: McClelland, Michael  
; APPLICANT: Welsh, John  
; APPLICANT: Trenkle, Thomas  
; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of  
; FILE REFERENCE: P-PH 3457  
; CURRENT APPLICATION NUMBER: US/09/300,958A  
; CURRENT FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/083,331  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/098,070  
; PRIOR FILING DATE: 1998-08-27  
; PRIOR APPLICATION NUMBER: 60/118,624  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 730  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-300-958A-16  
Alignment Scores:  
Pred. No.: 6,3e-52 Length: 730  
Score: 477.50 Matches: 87  
Percent Similarity: 72.34% Conservative: 49  
Best Local Similarity: 46.28% Mismatches: 51  
Query Match: 43.21% Indels: 1  
DB: 4 Gaps: 1  
US-09-817-198c-2 (1-212) x US-09-300-958A-16 (1-730)  
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20  
Db 82 ATGCGCCGGGACTACGACCACTCTTCAAGCTGCTCATCATCGCGCAGCGGTGGCG 141  
Qy 21 LysThrCysLeuLeuCysAtgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40  
Db 142 AAGAGCAGTTTACTGTGGGTTTTCGAGACAACTTCTCAGGACGCTATCATCCACG 201  
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
Db 202 ATCGAGTGGATTTCAAGATCCGACCGTGGAGATCAACGGGGAGAGGTGAACCTGAC 261  
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArg 80  
Db 262 ATCTGGGACACAGCGGGCGAGGCGCTTCGCGACCATCATCTCCAGCTATTATCGGGG 321

Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
Db 322 ACCCAGCGGCTATTGTGGTTTACAGCTCACCAGTCCGAGTCTCTTGTCAAGTCAAG 381  
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
Db 382 CGTGGCTTCACGAATCAACCAAGAACTGT--GATGATGTGTGCGCAATATTAGTGGT 438  
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140  
Db 439 AATAAGATGACGACCTGAGCGGAGGTGGTGAGACGAAGATGCTACAAATTCGCC 498  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 499 GGGCAGATGGCATCCAGTTGTCGACGACGCGCAAGGAGATGTCAACGTGGAAGAG 558  
Qy 161 SerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeu 180  
Db 559 ATGTTCAATGTCATCAGGAGTGGTCTCCGAGCAAAAGAAACAACCTGGCAAAACAG 618  
Qy 181 ArgMetArgAlaSerAsnGluLeu 188  
Db 619 CAGCAGCAACACAGACGATGTG 642  
RESULT 4  
US-09-016-434-1422  
; Sequence 1422, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1422:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 723 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: G550059  
US-09-016-434-1422  
Alignment Scores:









; Sequence 959, Application US/09620312D

QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyr 143  
Db 463 GATATGACGACGAAAGAGTATCCGTTCCGAAGACGAGACCTTGGCCAGGAGTAC 522  
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163  
Db 523 GGTGTTCCCTTCTCGGAGACGACGCGCAAGACTGGCATGAATGTG----- 567  
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183  
Db 568 -----GAGTTAGCCTTCTTGGCCATCGCCAAAGAA-----CTGAATACCGG 609  
QY 184 AlaSerAsnGlu 187  
Db 610 GCCGGGCATCAG 621

## RESULT 11

US-09-484-970B-142  
; Sequence 142, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PA-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; PRIORITY FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 142  
; LENGTH: 2612  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1  
US-09-484-970B-142

Alignment Scores:  
Pred. No.: 7,15e-45 Length: 2612  
Score: 429.00 Matches: 82  
Percent Similarity: 68.48% Conservative: 44  
Best Local Similarity: 44.57% Mismatches: 46  
Query Match: 38.82% Indels: 12  
DB: 4 Gaps: 3

## US-09-817-198C-2 (1-212) x US-09-484-970B-142 (1-2612)

QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24  
Db 91 TACGACCTCAGCGCAGGTGATGCTTCTGGAGACAGCGCTCGGCAACATGTTTC 150  
QY 25 LeuCysArgPheThrAspAsnGluPheHisSer---SerHisIleSerThrIleGlyVal 43  
Db 151 CTGATCCAAATTCAGAGCGGGCCCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATA 210  
QY 44 AspPheLysMetLysThrIleGluValAspGlyLysValArgIleGlnIleTyrAsp 63  
Db 211 GACTTCAGGAACAAGGTGGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 270  
QY 64 ThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGly 83  
Db 271 ACCGTGGGAGGACGGTTCGGAAGCGTCACCCATGCTTATTACAGAGTCTCAGGCC 330  
QY 84 IlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyrVal 103  
Db 331 TTGCTTCTGCTGTATGATACATCACCAACAATCTTCTTCGCAACAATCAGGCGCTGGCTC 390  
QY 104 SerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123  
Db 391 ACTGAGATTATGATATGCCAGAGGACGTCGTGTGATCATGCTGTAGGCAACAAGCGC 450

QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyr 143  
Db 451 GATATGACGACGAAAGAGTATCCGTTCCGAAGACGAGACCTTGGCCAGGAGTAC 510  
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163  
Db 511 GGTGTTCCCTTCTCGGAGACGACGCGCAAGACTGGCATGAATGTG----- 555  
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183  
Db 556 -----GAGTTAGCCTTCTTGGCCATCGCCAAAGAA-----CTGAATACCGG 597  
QY 184 AlaSerAsnGlu 187  
Db 598 GCCGGGCATCAG 609

## RESULT 12

US-09-620-312D-426  
; Sequence 426, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chonghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghaast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIORITY FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 426  
; LENGTH: 833  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (95)..(745)  
US-09-620-312D-426

Alignment Scores:  
Pred. No.: 5.68e-45 Length: 833  
Score: 424.00 Matches: 83  
Percent Similarity: 60.98% Conservative: 42  
Best Local Similarity: 40.49% Mismatches: 68  
Query Match: 38.37% Indels: 12  
DB: 4 Gaps: 2

## US-09-817-198C-2 (1-212) x US-09-620-312D-426 (1-833)

QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24  
Db 101 TATGCTTATCTTCAAGTATATCATCATCGGACACAGGTGTGGGAAGTCAATGCTC 160  
QY 25 LeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThrIleGlyValAsp 44







**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 08:59:39 ; Search time 311 Seconds  
(without alignments)  
2521.609 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAXQYDVLFRLLIGSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09817198/runat\_15032004\_101745\_20492/app\_query.fasta\_1.391  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomem62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09817198 @CIGN 1.1.333 @runat\_15032004\_101745\_20492  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubnpa/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubnpa/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
------------	-------	-------	-------	--------	-------	-------------

-----

1	1105	100.0	1054	12	US-10-363-616-74	Sequence 74, Appl
2	1105	100.0	2021	9	US-09-764-868-88	Sequence 88, Appl
3	1105	100.0	3257	9	US-09-817-198A-1	Sequence 1, Appl
4	832	75.3	566	9	US-09-764-868-507	Sequence 507, App
5	540	48.9	2411	15	US-10-264-049-290	Sequence 290, App
6	534.5	48.4	995	14	US-10-084-817-256	Sequence 256, App
7	534.5	48.4	1274	9	US-09-925-302-91	Sequence 91, Appl
8	530.5	48.0	1537	9	US-09-925-300-631	Sequence 631, Appl
9	530.5	48.0	3164	14	US-10-096-534-38	Sequence 38, Appl
10	528.5	47.8	624	9	US-09-794-257-9	Sequence 9, Appl
11	528.5	47.8	1161	9	US-09-794-257-7	Sequence 7, Appl
12	528.5	47.8	2497	9	US-09-834-975-879	Sequence 879, App
13	528.5	47.8	2497	9	US-09-834-975-885	Sequence 885, App
14	528.5	47.8	2497	9	US-09-834-975-896	Sequence 896, App
15	528.5	47.8	2497	9	US-09-834-975-896	Sequence 896, App
16	523	47.3	621	14	US-10-128-714-2241	Sequence 2241, Ap
17	523	47.3	621	14	US-10-128-714-7600	Sequence 7600, Ap
18	519	47.0	651	9	US-09-938-842A-836	Sequence 836, App
19	519	47.0	651	11	US-09-938-842A-836	Sequence 836, App
20	518	46.9	964	12	US-10-425-114-16115	Sequence 16115, A
21	514	46.5	1095	12	US-10-425-114-16115	Sequence 16115, A
22	513.5	46.5	1100	12	US-10-425-114-16320	Sequence 12044, A
23	513.5	46.5	1136	12	US-10-425-114-16320	Sequence 16320, A
24	513.5	46.5	1217	12	US-10-425-114-16320	Sequence 25934, A
25	513.5	46.5	1516	12	US-10-425-114-25934	Sequence 58714, A
26	512.5	46.4	880	12	US-10-425-114-24388	Sequence 24388, A
27	512.5	46.4	959	12	US-10-425-114-31413	Sequence 31413, A
28	512.5	46.4	1037	12	US-10-425-114-23752	Sequence 23752, A
29	512.5	46.4	1050	12	US-10-425-114-28854	Sequence 28854, A
30	511.5	46.3	1369	12	US-10-424-599-27513	Sequence 27513, A
31	507.5	45.9	1162	12	US-10-424-599-58679	Sequence 58679, A
32	506	45.8	1055	12	US-10-425-114-27209	Sequence 27209, A
33	483	43.7	745	14	US-10-128-714-1241	Sequence 1241, Ap
34	483	43.7	2745	14	US-10-128-714-241	Sequence 241, App
35	478.5	43.3	925	9	US-09-967-736-4	Sequence 4, Appl
36	478.5	43.3	939	14	US-10-102-806-48	Sequence 48, Appl
37	478.5	43.3	1898	12	US-10-221-278-5	Sequence 5, Appl
38	478.5	43.3	1898	15	US-10-291-172-5	Sequence 5, Appl
39	478.5	43.3	2078	14	US-10-119-428-5	Sequence 5, Appl
40	477.5	43.2	881	9	US-09-770-445-529	Sequence 529, App
41	477	43.2	1144	12	US-10-424-599-43925	Sequence 43925, A
42	476.5	43.1	879	12	US-10-425-114-20852	Sequence 20852, A
43	476	43.1	766	12	US-10-425-114-25791	Sequence 25791, A
44	476	43.1	915	12	US-10-425-114-2965	Sequence 2965, Ap
45	476	43.1	943	12	US-10-425-114-17495	Sequence 17495, A

ALIGNMENTS

RESULT 1  
US-10-363-616-74  
; Sequence 74, Application US/10363616  
; Publication No. US20040044181A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-113 (793)  
; CURRENT APPLICATION NUMBER: US/10363.616  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: 09/654,935  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 490  
; SEQ ID NO 74  
; LENGTH: 1054  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (81)..(719)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1054)  
; OTHER INFORMATION: n = a,t,c or g







Db 247 ATCCGAATTTCAAGATCCGCACTGTGGATATAGAGGGAAGAAATCAAACTACAA 306  
Qy 61 IletpaspThrAlaGlyGlnGluArgTyrglnThrileThrLysGlnTyrtYrArg 80  
Db 307 GTCTGGGACACGGTGGCCAGAGCGGTTCAAGACATTAATCTACTGCTACTACCGTGA 366  
Qy 81 AlaGlnGlyilePheLeuValTyrtAspIleSerSerGluArgSerTyrglnHisileMet 100  
Db 367 GCCATGGGCATTTATCTAGTATACACATCACCGATGAGAAATCTTTCGAGAATATTGAG 426  
Qy 101 LysTrpValSerAspValAspGluTyrtAlaProGluGlyValGlnLysileLeuileGly 120  
Db 427 AACTGGGATGAAAGCAATCAAGAGAAATGCCCTGGCTGGGTGGAGCGCTCTTCTGGGG 486  
Qy 121 AsnLysAlaAspGluGlnGlyArgGlnValGlyArgGlnGlyGlnGlnLeuAla 140  
Db 487 AACAAATGTGACATGGAGGCCAAGAGAGAGTGTGAGAGGAGCGCGGATAAGTTGGCT 546  
Qy 141 LysGluTyrglyMetAspPheTyrgluThrSerAlaCysThrAsnLeuAsnileLysGlu 160  
Db 547 CGAGAGCATGGAATCCGATTTTCGAACTAGTGTCTMAATCCAGTATGATGATGATGAG 606  
Qy 161 SerPheThrArgLeuThrGluLeuValLeu----- 170  
Db 607 GCTTTTAGTTCCTCCCTGGCCCGGACATCTTGTCTCAAGTCAGAGCGCGGAGATCAGGAAC 666  
Qy 171 -----GlnAlaHisArgLysGlu 176  
Db 667 GGCACAAAGCCTCCAGTACTGACCTGAAACTTGTGACAGAGAA-CACCAACAAGTG 725  
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196  
Db 726 CTCCTGGGCTGAGACCCCTTCTTGTCTCCCTCCCGGAGCTGAACCTGAGGGAGAC 785  
Qy 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211  
Db 786 AACGGCAGAGGAGTGTGACAGGGGAGAAATAGCAGAGGGGCTTGG 830  
RESULT 7  
US-09-925-302-91  
; Sequence 91, Application US/09925302  
; Patent No. US2002044941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 91  
; LENGTH: 1274  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1264)  
; NAME/KEY: misc feature  
; LOCATION: (1268)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-302-91  
Alignment Scores:  
Pred. No.: 4.1e-58 Length: 1274  
Score: 534.50 Matches: 102  
Percent Similarity: 66.38% Conservative: 54  
Best Local Similarity: 43.40% Mismatches: 55  
Query Match: 48.37% Indels: 25

DB: 9 Gaps: 1  
US-09-817-198C-2 (1-212) x US-09-925-302-91 (1-1274)  
Qy 1 MetAlaLysGlnTyrtAspValLeuPheArgLeuLeuileGlyAspSerGlyValGly 20  
Db 132 ATGCCCAAGCCCTACGACCACTCTTCAAGTGTCTGTGATCGGGGACTCGGGGGTGGGC 191  
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisileSerThr 40  
Db 192 AAGACTTGTCTGATCATCTGCTTTCAGAGGACAACCTTCAACACACTTACATCTCCACC 251  
Qy 41 IleGlyValAspPheLysMetLysThrileGluValAspGlyileLysValArgileGln 60  
Db 252 ATCGGAATTCATTTCAAGATCCGCACTGTGATATAGAGGGAAGAAATCAAACTACAA 311  
Qy 61 IletpaspThrAlaGlyGlnGluArgTyrglnThrileThrLysGlnTyrtYrArg 80  
Db 312 GTCTGGGACACGGTGGCCAGAGCGGTTCAAGACATTAATCTACTGCTACTACCGTGA 371  
Qy 81 AlaGlnGlyilePheLeuValTyrtAspIleSerSerGluArgSerTyrglnHisileMet 100  
Db 372 GCCATGGGCATTTATCTAGTATACACATCACCGATGAGAAATCTTTCGAGAATATTGAG 431  
Qy 101 LysTrpValSerAspValAspGluTyrtAlaProGluGlyValGlnLysileLeuileGly 120  
Db 432 AACTGGATGAAAGCAATCAAGAGATGCTCTGGCTGGGTGGAGCGCTCTTCTGGGG 491  
Qy 121 AsnLysAlaAspGluGlnGlyArgGlnValGlyArgGlnGlyGlnGlnLeuAla 140  
Db 492 AACAAATGTGACATGGAGGCCAAGAGAGAGTGTGAGAGGAGCGCGGATAAGTTGGCT 551  
Qy 141 LysGluTyrglyMetAspPheTyrgluThrSerAlaCysThrAsnLeuAsnileLysGlu 160  
Db 552 CGAGAGCATGGAATCCGATTTTCGAACTAGTGTCTMAATCCAGTATGATGATGATGAG 611  
Qy 161 SerPheThrArgLeuThrGluLeuValLeu----- 170  
Db 612 GCTTTTAGTTCCTCCCTGGCCCGGACATCTTGTCTCAAGTCAGAGCGCGGAGATCAGGAAC 671  
Qy 171 -----GlnAlaHisArgLysGlu 176  
Db 672 GGCACAAAGCCTCCAGTACTGACCTGAAACTTGTGACAGAGAA-CACCAACAAGTG 730  
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196  
Db 731 CTCCTGGGCTGAGACCCCTTCTTGTCTCCCGGAGCTGAACCTGAGGGAGAC 790  
Qy 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211  
Db 791 AACGGCAGAGGAGTGTGACAGGGGAGAAATAGCAGAGGGGCTTGG 835  
RESULT 8  
US-09-925-300-631  
; Sequence 631, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 631  
; LENGTH: 1537  
; TYPE: DNA  
; ORGANISM: Homo sapiens











**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:12:58 ; Search time 20 Seconds  
(without alignments)  
1019.630 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGKPGPANSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1077	97.5	212	F42148	GTP-binding protei
2	559	50.6	224	T233855	hypothetical prote
3	545	49.3	207	B49647	GTP-binding protei
4	545	49.3	207	B36364	GTP-binding protei
5	532	48.1	203	A49647	GTP-binding protei
6	531.5	48.1	200	A38625	GTP-binding protei
7	530.5	48.0	200	D36364	GTP-binding protei
8	530	48.0	209	B38625	GTP-binding protei
9	526	47.6	206	T78851	GTP-binding protei
10	522.5	47.3	200	B42148	GTP-binding protei
11	519	47.0	216	T45901	GTP-binding protei
12	514	46.5	201	T28971	hypothetical prote
13	512	46.3	216	S33900	GTP-binding protei
14	512	46.3	216	T48378	GTP-binding protei
15	510	46.2	200	S12790	GTP-binding protei
16	509	46.1	215	T14565	GTP-binding protei
17	508.5	46.0	215	S57478	GTP-binding protei
18	508	46.0	216	J50640	GTP-binding protei
19	507.5	45.9	222	T14405	small GTP-binding
20	506	45.8	216	S57471	GTP-binding protei
21	505.5	45.7	217	S36365	GTP-binding protei
22	502.5	45.5	215	S57462	GTP-binding protei
23	501	45.3	203	S51495	GTP-binding protei
24	495.5	44.8	204	JC7589	Sec4p homolog - Ye
25	492.5	44.6	215	S57474	GTP-binding protei
26	483.5	43.8	208	A34716	GTP-binding protei
27	483	43.7	208	A38202	GTP-binding protei
28	482	43.6	203	S34253	GTP-binding protei
29	480	43.4	202	S38740	GTP-binding protei

30	479	43.3	203	2	B34716	GTP-binding protei
31	478	43.3	202	2	S72515	GTP-binding protei
32	477.5	43.2	201	2	JC2488	GTP-binding protei
33	477.5	43.2	203	2	JC4105	GTP-binding protei
34	477	43.2	201	2	JC5337	GTP-binding protei
35	477	43.2	202	2	S41430	GTP-binding protei
36	476	43.1	258	2	B86153	ARA-5 (imported) -
37	475.5	43.0	203	2	JC1247	GTP-binding protei
38	475	43.0	203	2	S30096	GTP-binding protei
39	473.5	42.9	205	2	T33781	hypothetical prote
40	473	42.8	206	2	T14391	GTP-binding protei
41	473	42.8	210	2	T18242	ras protein homolo
42	470.5	42.6	203	2	B38202	GTP-binding protei
43	470.5	42.6	215	1	TVBQ04	GTP-binding protei
44	469	42.4	205	2	S38339	GTP-binding protei
45	468	42.4	205	1	TVRTVP	GTP-binding protei

ALIGNMENTS

RESULT 1

F42148  
GTP-binding protein rab15 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 02-Feb-2001  
C:Accession: F42148  
R:Elferink, L.A.; Anzai, K.; Scheller, R.H.  
J. Biol. Chem. 267, 5768-5775, 1992  
A:Title: rab15, a novel low molecular weight GTP-binding protein specifically expressed i  
A:Reference number: A42148; MUID:92210533; PMID:1313420  
A:Accession: F42148  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-212 <HELP>  
A:Cross-references: GB:M83679; NID:G206536; PIDN:AAA41995.1; PID:G206537  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine  
F:9-124/Domain: translation elongation factor Tu homology <ETU>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:210,212/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match	97.5%	Score 1077;	DB 2;	Length 212;
Best Local Similarity	97.6%	Pred. No. 1.2e-78;		
Matches	207;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0;
Qy	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ	60	
Db	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ	60	
Qy	61	IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILIG	120	
Db	61	IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILIG	120	
Qy	121	NKADEEQKRVGRQGOQLAKEYGMDFYETSACTNLNINKESFTLRLTELVLQAHKRELEGL	180	
Db	121	NKADEEQKRVGRQGOQLAKEYGMDFYETSACTNLNINKESFTLRLTELVLQAHKRELDGL	180	
Qy	181	RMASNELALAELEEEKGKPGPANSKTCWC	212	
Db	181	RTCASNELALAELEDEGKTEGPANSKTCWC	212	

RESULT 2

T33855  
hypothetical protein D1037.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 11-Jan-2000  
C:Accession: T33855  
R:Ledwith, J.; Biewald, T.  
Submitted to the EMBL Data Library, November 1998  
A:Description: The sequence of C. elegans cosmid D1037.

A;Reference number: Z21424  
A;Accession: T33855  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-224 <LEAD>  
A;Cross-references: EMBL:AF106592; PIDN:AACT8494.1; GSPDB:GN00019; CESP:D1037.4  
A;Experimental source: strain Bristol N2; clone D1037.  
C;Genetics:  
A;Gene: CESP:D1037.4  
A;Map position: 1  
A;Introns: 10/3; 62/2; 82/3; 181/1  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 50.6%; Score 559; DB 2; Length 224;  
Best Local Similarity 55.3%; Pred. No. 2.4e-37;  
Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNNSPISTIGIDFKIRTIELDGKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Qy 181 RMRASNEAL 190  
Db 177 EMRAATGAAI 186

## RESULT 3

B49647  
GTP-binding protein rab8 - human  
C;Species: Homo sapiens (man)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Jan-2001  
R;Accession: B49647; S36817  
C;Accession: B49647; S36817  
J. Cell Biol. 124, 101-115, 1994  
A;Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cell  
A;Reference number: A49647; MUID:94124602; PMID:8294494  
A;Accession: B49647  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-207 <ZAH>  
A;Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318  
R;Joberty, G.; Tavittian, A.; Zahraoui, A.  
FEBS Lett. 330, 323-328, 1993  
A;Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.  
A;Reference number: S36817; MUID:93387463; PMID:8375503  
A;Accession: S36817  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 175-186 <JOB>  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding SAK/L motif  
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.8e-36;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNNSPISTIGIDFKIRTIELDGKIKLQ 60

Qy 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Qy 181 RMRASNE 187  
Db 181 SPOGSNQ 187

## RESULT 4

B36364  
GTP-binding protein rab8 - dog  
C;Species: Canis lupus familiaris (dog)  
C;Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001  
C;Accession: B36364; S15604  
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.  
Mol. Cell. Biol. 10, 6578-6585, 1990

A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.  
A;Reference number: A36364; MUID:91061765; PMID:2123294  
A;Accession: B36364  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-207 <CHA>  
A;Cross-references: GB:X56385; NID:g920; PIDN:CAB56776.1; PID:g6006436  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding SAK/L motif  
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.8e-36;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNNSPISTIGIDFKIRTIELDGKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAGQIFLVYDTSERSYOHIMKWSDVDVEYAPGQVKILIG 120  
Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Qy 181 RMRASNE 187  
Db 181 SPOGSNQ 187

## RESULT 5

A49647  
GTP-binding protein Rabi3 - human  
C;Species: Homo sapiens (man)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Feb-2001  
C;Accession: A49647

R;Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavittian, A.; Louvard  
J. Cell Biol. 124, 101-115, 1994  
A;Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cells  
A;Reference number: A49647; MUID:94124602; PMID:8294494  
A;Accession: A49647  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-203 <ZAH>  
A;Cross-references: EMBL:X75593; NID:g452319; PIDN:CAA53266.1; PID:g452320  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding SAK/L motif  
F;212,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta  
F;200/Binding site: farnesyl (Cys) (covalent) #status predicted

Query Match 48.1%; Score 532; DB 2; Length 203;  
Best Local Similarity 47.6%; Pred. No. 3e-35;  
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRIQ 60  
Db 1 MAKAYDLFLKLLIGDSGVGKTCIIIRFAEDFNNTVISTIGIDFKIRTVDIEGKKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGQKILIG 120  
Db 61 VWDTAGQERFHTITAYYRGAMGIMLVYDITDEKSFENIQNMWKSIRENASAGVERLLLG 120  
Qy 121 NKADDEOKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKELEGL 180  
Db 121 NKCDMEAKRVQKQADKLAREHGIRFFETSAKSMNVDEAFSLARDIL-----LKSG 174  
Qy 181 RMASNELALAELEEKGKPGPANSKTC 210  
Db 175 GRRSGN-----GNKP--PSTD LKTC 192

## RESULT 6

A38625  
GTP-binding protein oral - electric ray (Discopyge ommata)  
C;Species: Discopyge ommata  
C;Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 02-Feb-2001  
C;Accession: A38625  
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.  
J. Biol. Chem. 266, 2675-2680, 1991  
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.  
A;Reference number: A38625; MUID:91115900; PMID:1899244  
A;Accession: A38625  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-200 <NGS>  
A;Cross-references: GB:M38390; NID:g213114; PIDN:AAA49230.1; PID:g213115  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop  
F;10-125/Domain: translation elongation factor Tu homology <ETU>  
F;16-23/Region: nucleotide-binding motif A (P-loop)  
F;122-125/Region: GTP-binding NKXD motif  
F;152-154/Region: GTP-binding SAK/L motif

Query Match 48.1%; Score 531.5; DB 2; Length 200;  
Best Local Similarity 56.4%; Pred. No. 3.2e-35;  
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRI 59  
Db 1 MAKTYDVLFLKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVELGKKIKL 60  
Qy 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGQKILI 119  
Db 61 QIWDTAGQERFHTITAYYRGAMGIMLVYDITNAKSFENISKWLNRIDEHANEDVERMLL 120  
Qy 120 NKADDEOKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQ 171  
Db 121 GNKCDMEDKRVLSKGRQIAREHAIRFPFETSAKANINIEKAFLLTAEIDILQ 172

## RESULT 7

D36364  
GTP-binding protein rab10 - dog  
C;Species: Canis lupus familiaris (dog)  
C;Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001

C;Accession: D36364; S15601  
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.  
Mol. Cell. Biol. 10, 6578-6585, 1990  
A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.  
A;Reference number: A36364; MUID:91061765; PMID:2123294  
A;Accession: D36364  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-200 <CHA>  
A;Cross-references: GB:X56387; NID:g914; PIDN:CAA39798.1; PID:g915  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;10-125/Domain: translation elongation factor Tu homology <ETU>  
F;16-23/Region: nucleotide-binding motif A (P-loop)  
F;122-125/Region: GTP-binding NKXD motif  
F;152-154/Region: GTP-binding SAK/L motif  
F;199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530.5; DB 2; Length 200;  
Best Local Similarity 55.8%; Pred. No. 3.8e-35;  
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRI 59  
Db 1 MAKTYDVLFLKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVELGKKIKL 60  
Qy 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGQKILI 119  
Db 61 QIWDTAGQERFHTITAYYRGAMGIMLVYDITNGKSFENISKWLNRIDEHANEDVERMLL 120  
Qy 120 GNKADDEOKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQ 171  
Db 121 GNKCDMDKRVLPKGEQIAREHGIRFFETSAKVNINIEKAFLLTAEIDILR 172

## RESULT 8

B38625  
GTP-binding protein ora2 - electric ray (Discopyge ommata)  
C;Species: Discopyge ommata  
C;Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 02-Feb-2001  
C;Accession: B38625  
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.  
J. Biol. Chem. 266, 2675-2680, 1991  
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.  
A;Reference number: A38625; MUID:91115900; PMID:1899244  
A;Accession: B38625  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-209 <NGS>  
A;Cross-references: GB:M38391  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;150-152/Region: GTP-binding SAK/L motif  
F;206/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530; DB 2; Length 209;  
Best Local Similarity 47.4%; Pred. No. 4.4e-35;  
Matches 99; Conservative 56; Mismatches 44; Indels 10; Gaps 3;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRIQ 60  
Db 1 MAKTYDVLFLKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVELGKKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGQKILIG 120  
Db 61 IWDTAGQERFHTITAYYRGAMGIMLVYDITNEKSFENIKWLNRIDEHANEDVERMLG 120  
Qy 121 NKADDEOKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKELEGL 180  
Db 121 NKCDMNEKRVSKERGEKLAIDYGIKF-ETSAKSSINVEEAFITLARDIMTKLNKQM--- 176



A;Residues: 1-201 <WAM>  
A;Cross-references: EMBL:U80033; PIDN:AA48200.1; GSPDB:GN00019; CESP:T23H2.5  
A;Experimental source: strain Bristol N2; clone T23H2  
C;Genetics:  
A;Gene: CESP:T23H2.5  
A;Map position: 1  
A;Introns: 43/1; 147/2; 173/3  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 46.5%; Score 514; DB 2; Length 201;  
Best Local Similarity 46.9%; Pred. No. 8e-34;  
Matches 100; Conservative 48; Mismatches 51; Indels 14; Gaps 3;

Qy 1 MAKQ-YDVLFRLLIGSGVGKTCLLCRFTDNEFHSHSTIGVDYFQKMTIEVDGKVRKI 59  
Db 1 MARRPYDMLFKLLIGSGVGKTCILYRSDDAFTTFISTIGIDFKIKTIELKGGKIKL 60

Qy 60 QIWDTAGQERYQTITKQYVRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILI 119  
Db 61 QIWDTAGQERFHTITTSYIRGAMGIMLVYDITNAKSFDTNAKWLRIDEHASEDVVKMIL 120

Qy 120 GNKADERQKRVGREGQOLAKYGMDFYETSACTNLNIKESFTRLTELVLQAHKKELEG 179  
Db 121 GNKCDMSDRVVSREGEKIAQDGHGISHFETSAKLNHVDTAFYDLAEAIL----- 171

Qy 180 LRMRASNELALAELEEBEGKEGPEGPANSSKTCWC 212  
Db 172 AKMPDSTD-----EQSRDVTNPVQVQROSSGGC 200

RESULT 13  
S33900  
GTP-binding protein ypt2 - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Feb-2001  
C;Accession: S33900; JQ2233  
R;Fleming, A.J.; Mandel, T.; Roth, I.; Kuhlmeier, C.  
Plant Cell 5, 297-309, 1993  
A;Title: The patterns of gene expression in the tomato shoot apical meristem.  
A;Reference number: S33899; MUID:93222691; PMID:8467223  
A;Accession: S33900  
A;Molecule type: mRNA  
A;Residues: 1-216 <FL2>  
A;Cross-references: GB:X69980; NID:g313028; PIDN:CAA49600.1; PID:g313029  
C;Genetics:  
A;Gene: ypt2  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; membrane protein; nucleotide binding; P-loop  
F;16-131/Domain: translation elongation factor Tu homology <ETU>  
F;22-29/Region: nucleotide-binding motif A (P-loop)  
F;128-131/Region: GTP-binding NKXD motif  
F;159-161/Region: GTP-binding SAK/L motif

Query Match 46.3%; Score 512; DB 2; Length 216;  
Best Local Similarity 50.5%; Pred. No. 1.3e-33;  
Matches 99; Conservative 41; Mismatches 44; Indels 12; Gaps 3;

Qy 5 YDVLFRLLIGSGVGKTCILCRFTDNEFHSHSTIGVDYFQKMTIEVDGKVRKIOWDT 64  
Db 12 YDYLKLLIGDGTGVGKSCLLRFSDDFTFTTIGIDFKIRTIELDGKRIKLQIOWDT 71

Qy 65 AGQERYOTITKQYVRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILGNKAD 124  
Db 72 AGQERFRTITAYYRGAMGILLVYDVTDESFNIRNWKNIHQHASNVDNKNILVGNKAD 131

Qy 125 -EEQKRVGREGQOLAKYGMDFYETSACTNLNIKESF-----TRLTELVLQAHK 175  
Db 132 MDESKRAVPTSKGQALADEYGIKPFETSAKTNLNVENVFMSIAKDQKRLTETDTKA--- 188

Qy 176 ELEGLRMRASNELALAELEE 191  
Db 189 EPQIRINQSDQAGTA 204

RESULT 14  
T48378  
GTP-binding protein-like - Arabidopsis thaliana  
N;Alternate names: protein F12E4.300  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000  
C;Accession: T48378  
R;Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24492  
A;Accession: T48378  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-216 <BEV>  
A;Cross-references: EMBL:AL162751  
A;Experimental source: cultivar Columbia; BAC clone F12E4  
C;Genetics:  
A;Map position: 5  
A;Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3  
A;Note: F12E4.300  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
F;16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 512; DB 2; Length 216;  
Best Local Similarity 49.8%; Pred. No. 1.3e-33;  
Matches 100; Conservative 43; Mismatches 46; Indels 12; Gaps 3;

Qy 5 YDVLFRLLIGSGVGKTCILCRFTDNEFHSHSTIGVDYFQKMTIEVDGKVRKIOWDT 64  
Db 12 YDYLKLLIGDGTGVGKSCLLRFSDDFTFTTIGIDFKIRTIELDGKRIKLQIOWDT 71

Qy 65 AGQERYOTITKQYVRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILGNKAD 124  
Db 72 AGQERFRTITAYYRGAMGILLVYDVTDESFNIRNWKNIHQHASNVDNKNILVGNKAD 131

Qy 125 -EEQKRVGREGQOLAKYGMDFYETSACTNLNIKESF-----TRLTELVLQAHK 175  
Db 132 MDESKRAVPTAKGQALADEYGIKPFETSAKTNLNVENVFMSIAKDQKRLTETDTKA--- 188

Qy 176 ELEGLRMRASNELALAELEE 196  
Db 189 EPOGKITQDTAAASSSTAER 209

RESULT 15  
S12790  
GTP-binding protein ypt2 - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 02-Feb-2001  
C;Accession: S12790; S10493; T39214  
R;Haubruck, H.; Engelke, U.; Mertins, P.; Gallwitz, D.  
EMBO J. 9, 1957-1962, 1990  
A;Title: Structural and functional analysis of ypt2, an essential ras-related gene in the  
fission yeast Schizosaccharomyces pombe.  
A;Reference number: S12790; MUID:90269232; PMID:2112089  
A;Accession: S12790  
A;Molecule type: DNA  
A;Residues: 1-200 <HAU>  
A;Cross-references: EMBL:X52469; NID:g5143; PIDN:CAA36707.1; PID:g5144  
R;Fawell, E.; Hook, S.; Sweet, D.; Armstrong, J.  
Nucleic Acids Res. 18, 4264, 1990  
A;Title: Novel YPT1-related genes from Schizosaccharomyces pombe.  
A;Reference number: S10492; MUID:903332438; PMID:2115995  
A;Accession: S10493  
A;Molecule type: DNA  
A;Residues: 1-200 <FAW>  
A;Cross-references: EMBL:X52864; NID:g5149; PIDN:CAA37045.1; PID:g5150  
R;McDougall, R.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A;Reference number: Z21836  
A;Accession: T39214  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

```

A;Residues: 1-200 <MCD>
A;Cross-references: EMBL:Z99262; PIDN:CA816405.1; GSPDB:GN000666; SPDB:SPAC39E9.07C
A;Experimental source: strain 972h-; cosmid c9E9
C;Genetics:
A;Gene: ypt2
A;Map position: 1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;10-125/Domain: translation elongation factor Tu homology <TU>
F;16-23/Region: nucleotide-binding motif A (P-loop)
F;122-125/Region: GTP-binding NKXD motif
F;152-154/Region: GTP-binding SAK/L motif
F;199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match          46.2%; Score 510; DB 2; Length 200;
Best Local Similarity 51.9%; Pred. No. 1.6e-33;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KQYDVLFRLLIGDSGVGKTCLLCRFTDNEPSSHISTIGVDYFKMKTIEVDGKVRIGIW 62
   |||||
Db 4 KSYDYLIKALLIGDSGVGKSCLLRFSEDSFSPFITIGIDFKIRTELGGKIKLQIW 63
   |||||

QY 63 DTAGQERYQITIKQYRRAGQIFVLVYDISSRSYQHIMKWVSDVDEYAPGQVKILGNK 122
   |||||
Db 64 DTAGQERFRITATYYRGAMGILLLYDVTDKKSPDNVTRTFPSNVEQHASENVYKILGNK 123
   |||||

QY 123 ADREQKQVGRQEQOQLAKEYGMDFYETSACTNINIKESFTRLTELVLQAHRKELEGIRM 182
   |||||
Db 124 CDCEQKQVSEFQEQQALADELGVKFLASAKTNVNDFAFTLA-----REIKKQKI 175
   |||||

QY 183 RASNELA 189
   |||||
Db 176 DAENEPS 182

```

Search completed: March 15, 2004, 11:16:52  
Job time : 21 secs





```
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAN; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT LIPID 210 210 S-geranylgeranyl cysteine
FT (By similarity).
FT LIPID 212 212 S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 212 AA; 24390 MW; D16A0C71797ED782 CRC64;

Query Match 100.0%; Score 1105; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 5e-81;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60

Qy 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSDVDVEYAPGVQKILIG 120
Db 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSDVDVEYAPGVQKILIG 120

Qy 121 NKADEQKRVGREGQOQLAKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180
Db 121 NKADEQKRVGREGQOQLAKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180

Qy 181 RMASNELALAELEEEGKPGPANSSKTCWC 212
Db 181 RMASNELALAELEEEGKPGPANSSKTCWC 212

RESULT 2
RB15 MOUSE STANDARD; PRT; 212 AA.
AC Q8K386;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-15.
GN RAB15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;
```

```
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May act in concert with RAB1A in regulating aspects of
CC synaptic vesicle membrane flow within the nerve terminal (By
CC similarity).
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC027769; AH27769.1; -.
DR MGD; MGI:1916865; 2310012G06Rik.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR003577; GTPase_Ras.
DR InterPro; IPR003578; GTPase_Rho.
DR InterPro; IPR002041; RAN.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAN; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT LIPID 210 210 S-geranylgeranyl cysteine
FT (By similarity).
FT LIPID 212 212 S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 212 AA; 24318 MW; ESC492846DD47F12 CRC64;

Query Match 98.6%; Score 1090; DB 1; Length 212;
Best Local Similarity 98.1%; Pred. No. 7.9e-80;
Matches 208; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60

Qy 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSDVDVEYAPGVQKILIG 120
Db 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYOHIMKWSDVDVEYAPGVQKILIG 120

Qy 121 NKADEQKRVGREGQOQLAKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180
Db 121 NKADEQKRVGREGQOQLAKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180

Qy 181 RMASNELALAELEEEGKPGPANSSKTCWC 212
Db 181 RMASNELALAELEEEGKPGPANSSKTCWC 212

RESULT 3
RB15 RAT STANDARD; PRT; 212 AA.
AC P35289;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-15.
GN RAB15.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=92210533; PubMed=1313420;  
 RX Eiferink L.A., Anzai K., Scheller R.H.;  
 RT "Rab15, a novel low molecular weight GTP-binding protein specifically  
 RT expressed in rat brain.";  
 RL J. Biol. Chem. 267:5768-5775 (1992).  
 RN [2]  
 RN ERRATUM.  
 RX MEDLINE=93054572; PubMed=1429617;  
 RA Eiferink L.A., Anzai K., Scheller R.H.;  
 RL J. Biol. Chem. 267:22693-22693 (1992).  
 CC -!- FUNCTION: May act in concert with RAB3A in regulating aspects of  
 CC synaptic vesicle membrane flow within the nerve terminal.  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in neural tissues.  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M83679; AAA41995.1; -.  
 DR PIR; F42148; F42148.  
 DR HSPSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF000071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFPAMS; TIGR00231; small GTP; 1.  
 DR GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT LIPID 210 210 S-geranylgeranyl cysteine  
 FT (By similarity).  
 FT LIPID 212 212 S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 212 AA; 24283 MW; 04817DDA66CADE12 CRC64;  
 Query Match 97.5%; Score 1077; DB 1; Length 212;  
 Best Local Similarity 97.6%; Pred. No. 8.5e-79;  
 Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRPTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60  
 DB 1 MAKQYDVLFRLLIGDSGVGKTCLLCRPTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60  
 QY 61 IWDTAGQERYTITKQYRRAQGIPLVYDIDSSERSYCHMKWSDVDVEYAPGVQKILIG 120  
 DB 61 IWDTAGQERYTITKQYRRAQGIPLVYDIDSSERSYCHMKWSDVDVEYAPGVQKILIG 120  
 QY 121 NKADBEOKRVGREGQOQALAKYGMDFVETSACTNLNLIKESFTRLTLVLAQHRKELEGL 180  
 DB 121 NKADBEOKRVGREGQOQALAKYGMDFVETSACTNLNLIKESFTRLTLVLAQHRKELEGL 180  
 QY 181 RMRASNELALAELEBEKGKPEGPANSSKTCWC 212  
 DB 181 RTCASNELALAELEDEKGTEGPANSSKTCWC 212  
 RESULT 4  
 RAB8 HUMAN  
 ID -RAB8 HUMAN  
 AC P24407;

DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ras-related protein Rab-8 (Rab-8A) (Oncogene c-mel).  
 GN MEL OR RAB8.  
 GN Homo sapiens (Human), and  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606, 9615;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX MEDLINE=94124602; PubMed=8294494;  
 RX Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,  
 RA Tavittian A., Louvard D.;  
 RT "A small rab GTPase is distributed in cytoplasmic vesicles in non  
 RT polarized cells but colocalizes with the tight junction marker ZO-1  
 RT in polarized epithelial cells.";  
 RL J. Cell Biol. 124:101-115 (1994).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX MEDLINE=91360267; PubMed=1886711;  
 RX Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
 RA Johnson K.J.;  
 RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
 RT genes.";  
 RL Oncogene 6:1347-1351 (1991).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Brain;  
 RX PubH H.L. III, Ikeda S.R., Aronstam R.S.;  
 RT "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guthrie cDNA resource center ([www.cdna.org](http://www.cdna.org)).";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Skin;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=C.familiaris; STRAIN=Cocker spaniel;  
 RX MEDLINE=91061765; PubMed=2123294;  
 RX Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;  
 RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell  
 RT line.";  
 RL Mol. Cell Biol. 10:6578-6585 (1990).  
 CC -!- FUNCTION: May be involved in vesicular trafficking and  
 CC neurotransmitter release.  
 CC -!- SUBUNIT: Interacts with MAP4K2 and SYT4 (By similarity).  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 DR EMBL; X56741; CA40065.1; --  
 DR EMBL; S53268; AAB19881.1; --  
 DR EMBL; AF498943; AAM21091.1; --  
 DR EMBL; BC002977; AAH02977.1; --  
 DR EMBL; X56385; CAB56776.1; --  
 DR PIR; B36364; B36364.  
 DR PIR; B49647; B49647.  
 DR HSP; P05713; 3RAB.  
 DR Genew; HGNC:7007; MEL.  
 DR MIN; 165040; --  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMS; TIGR00231; small GTP; 1.  
 KW GTP-binding; Prenylation; Lipoprotein; Protein transport;  
 KW Proto-oncogene.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 204 204 S-geranylgeranyl cysteine  
 FT (BY similarity).  
 FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).  
 SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD056 CRC64;

Query Match 49.3%; Score 545; DB 1; Length 207;  
 Best Local Similarity 52.9%; Pred. No. 1.6e-36;  
 Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIVRIQ 60  
 DB 1 MAKTYDYLKLLIGDSGVGKTCCLFRFSDAFNSTFISTIGIDFKIRTIELDGKRIKQ 60  
 QY 61 IWDTAGQERYQTITKQYRAAQGIFLVYDTSERSYQHIMKWSVDVDEYAPGVOKILIG 120  
 DB 61 IWDTAGQERYQTITTAIRGAMGIMLVYDITNEKSFNIRNWIKNIEHASSADVVERMILG 120  
 QY 121 NKADQKQVGRGQOQLAKEYGMDFYETSACTNLNIKESFTRLTLVLQAHKKELEGL 180  
 DB 121 NKCDVNDKQVSKERGEKLAIDYGIKFWETSAKANINVENAFFTLARDIKAKMDKKLEGN 180  
 QY 181 RMRASNE 187  
 DB 181 SPOGSNQ 187

RESULT 5  
 RAB8 DISOM STANDARD; PRT; 210 AA.  
 ID -RAB8 DISOM STANDARD; PRT; 210 AA.  
 AC P22128;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein Rab-8 (ORA2).  
 OS Discopryge omata (Electric ray).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Hypnosqualea; Pristiorajae; Batoidae;  
 OC Torpediniformes; Narcinoidei; Narcinidae; Discopryge.  
 OX NCBI\_TaxID=7785;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Electric lobe;

RX MEDLINE=91115900; PubMed=1899244;  
 RA Ngeese J.K., Elferink L.A., Scheller R.H.;  
 RT "A family of ras-like GTP-binding proteins expressed in electromotor  
 RT neurons.";  
 RL J. Biol. Chem. 266:2675-2680 (1991).  
 CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 DR EMBL; M38391; AAA49232.1; --  
 DR PIR; B38625; B38625.  
 DR HSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMS; TIGR00231; small GTP; 1.  
 KW GTP-binding; Prenylation; Lipoprotein.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 207 207 S-geranylgeranyl cysteine  
 FT (BY similarity).  
 SQ SEQUENCE 210 AA; 24164 MW; 0F73EDB0DB8B9EEA CRC64;

Query Match 48.8%; Score 539.5; DB 1; Length 210;  
 Best Local Similarity 47.4%; Pred. No. 4.5e-36;  
 Matches 99; Conservative 56; Mismatches 45; Indels 9; Gaps 2;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIVRIQ 60  
 DB 1 MAKTYDYLKLLIGDSGVGKTCCLFRFSDAFNSTFISTIGIDFKIRTIELDGKRIKQ 60  
 QY 61 IWDTAGQERYQTITKQYRAAQGIFLVYDTSERSYQHIMKWSVDVDEYAPGVOKILIG 120  
 DB 61 IWDTAGQERYQTITTAIRGAMGIMLVYDITNEKSFNIRNWIKNIEHASSADVVERMILG 120  
 QY 121 NKADQKQVGRGQOQLAKEYGMDFYETSACTNLNIKESFTRLTLVLQAHKKELEGL 180  
 DB 121 NKCDVNDKQVSKERGEKLAIDYGIKFWETSAKANINVENAFFTLARDIKAKMDKKLEGN 180  
 QY 181 RMRASNELALAE-LEEEGKPEGPANSSK 208  
 DB 178 -----NENSLQEAVDKLPKPKPKSKK 201

RESULT 6  
 RB13 HUMAN STANDARD; PRT; 203 AA.  
 ID -RB13 HUMAN STANDARD; PRT; 203 AA.  
 AC P51153;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ras-related protein Rab-13.  
 GN RAB13.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94124602; PubMed=8294494;  
 RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,  
 RA Tavittian A., Louvard D.;

RT "A small rab GTPase is distributed in cytoplasmic vesicles in non  
 RT polarized cells but colocalizes with the tight junction marker ZO-1  
 RT in polarized epithelial cells.";  
 RL J. Cell Biol. 124:101-115(1994).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RC Puhl H.L. III, Ikeda S.R., Aronstam R.S.;  
 RA "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Placenta;  
 RC MEDLINE=42388257; PubMed=12477932;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Could participate in polarized transport, in the  
 CC assembly and/or the activity of tight junctions.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC TIGHT JUNCTIONS OR ASSOCIATED  
 CC WITH VESICLES SCATTERED THROUGHOUT THE CYTOPLASM IN CELLS LACKING  
 CC TIGHT JUNCTIONS.  
 CC -!- TISSUE SPECIFICITY: Detected in several types of epithelia,  
 CC including intestine, kidney, liver, and in endothelial cells.  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75593; CAA53266.1; -.  
 DR EMBL; AF498948; AAM21096.1; -.  
 DR EMBL; BC000799; AAH00799.1; -.  
 DR PIR; A49647; A49647.  
 DR HSSP; P01112; 1PLI.  
 DR Genew; HGNC:9762; RAB13.  
 DR MIM; 602672; -.  
 DR GO; GO:0005923; C:tight junction; TAS.  
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 DR GO; GO:0016192; P:vesicle-mediated transport; TAS.  
 DR InterPro; IPR003579; GTPase\_Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR GTP-binding; Lipoprotein; Prenylation.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT LIPID

FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFECTOR REGION (BY SIMILARITY).  
 FT LIPID 200 200 S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 203 AA; 22774 MW; 141621CB998178DA CRC64;  
 Query Match 48.1%; Score 532; DB 1; Length 203;  
 Best local similarity 47.6%; Pred. No. 1.7e-35;  
 Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;  
 QY 1 MAKQYDVLFRLLILIGDSGVKTCCLLCPTDNEHSSHSITGVDFKMTIEVDGKVRIO 60  
 DB 1 MAKAYDHLFKLLILIGDSGVKTCCLIRPAEDNFNTYITIGIDFKIRTVDIEGKKIKQ 60  
 QY 61 IWTAGQRYQTITKQYRRAQGIYVDITSSERSYQHIMKWSVDVDEYAPGVQKILIG 120  
 DB 61 VWTAGQERKTIITAYRGAMGILLVYDITDEKSFENIQNWMSIKENASAGVERLLLG 120  
 QY 121 NKADEEOKRVGREQGOOLAKYGMDFYTSACTNLNKSFTLTLVLAQHRKEGL 180  
 DB 121 NKCDMEAKRKVQEQADKLAREHGIRFETTSKSSMMVDEAFSLARDIL-----LKSG 174  
 QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210  
 DB 175 GRRSGN-----GNKP--PSTDLKTC 192

RESULT 7  
 RB10\_DISOM STANDARD; PRT; 200 AA.  
 AC P22127;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein Rab-10 (ORA1).  
 OS Discopoge ommata (Electric ray).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Hypnosquala; Pristiogadidae; Batoidae;  
 OC Torpediniformes; Narcinoidei; Narcinidae; Discopoge.  
 OX NCBI\_TaxID=7785;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Electric lobe;  
 RX MEDLINE=91115900; PubMed=1899244;  
 RA Ngeee J.K., Elferink L.A., Scheller R.H.;  
 RT "A family of ras-like GTP-binding proteins expressed in electromotor  
 RT neurons.";  
 RL J. Biol. Chem. 266:2675-2680(1991).  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M38390; AAA49230.1; -.  
 DR PIR; A38625; A38625.  
 DR HSSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR GTP-binding; Lipoprotein; Prenylation.  
 KW GTP-binding; Lipoprotein; Prenylation.  
 FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
 FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
 FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
 FT LIPID 198 198 S-geranylgeranyl cysteine

FT LIPID 199 199 (By similarity).  
 FT S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 200 AA; 22623 MW; 41D38E3D760519C5 CRC64;  
 Query Match 48.1%; Score 531.5; DB 1; Length 200;  
 Best Local Similarity 56.4%; Pred. No. 1.8e-35;  
 Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;  
 QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIKVRI 59  
 Db 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIKVRI 60  
 QY 60 QIWDTAGQERYQITTKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119  
 Db 61 QIWDTAGQERFHTITTSYYRGAMIMLVYDITNAKSPENISKLNRNIDEHANEDEVRLML 120  
 QY 120 GNKADBEQKQVQREOQQLAKYGMDFYETSACTNLNLIKESFTRLTLEVLQ 171  
 Db 121 GNKCDMEDKRVVLKSKGEQIAREHAIRFFETSAKANINIEKAFITLAEIDLQ 172  
 RESULT 8  
 RB10 CANFA STANDARD; PRT; 200 AA.  
 AC P24409;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein Rab-10.  
 GN RAB10.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Cocker spaniel;  
 RX MEDLINE=91061765; PubMed=2123294;  
 RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;  
 RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell  
 line.";  
 RL Mol. Cell. Biol. 10:6578-6585(1990).  
 CC -1- FUNCTION: May be involved in vesicular trafficking and  
 CC neurotransmitter release.  
 CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X56387; CAA39798.1; ..  
 DR PIR; D36364; D36364.  
 DR HSRP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRfams; TIGR00231; small\_GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
 FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
 FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
 FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 199 199 S-geranylgeranyl cysteine  
 FT (By similarity).  
 FT S-geranylgeranyl cysteine

FT SQ SEQUENCE 200 AA; 22569 MW; 5D52B8E8E47D4362 CRC64;  
 (By similarity).  
 Query Match 48.0%; Score 530.5; DB 1; Length 200;  
 Best Local Similarity 55.8%; Pred. No. 2.2e-35;  
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;  
 QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIKVRI 59  
 Db 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIKVRI 60  
 QY 60 QIWDTAGQERYQITTKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119  
 Db 61 QIWDTAGQERFHTITTSYYRGAMIMLVYDITNGKSFENISKLNRNIDEHANEDEVRLML 120  
 QY 120 GNKADBEQKQVQREOQQLAKYGMDFYETSACTNLNLIKESFTRLTLEVLQ 171  
 Db 121 GNKCDMEDKRVVLKSKGEQIAREHAIRFFETSAKANINIEKAFITLAEIDL 172  
 RESULT 9  
 RB10 HUMAN STANDARD; PRT; 200 AA.  
 AC O88386; Q9D7X6;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ras-related protein Rab-10.  
 GN RAB10.  
 OS Homo sapiens (Human), and  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606, 10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Pituitary;  
 RX MEDLINE=20402571; PubMed=10931946;  
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
 RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
 RT axis and full-length cDNA cloning.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX Wong K., Hong W., Tang B.;  
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wakatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Brain;  
 RX Puhl H.L. III, Ikeda S.R., Aronstam R.S.;  
 RA "cDNA clones of human proteins involved in signal transduction  
 RA sequenced by the Guthrie cDNA resource center ([www.cdna.org](http://www.cdna.org)).";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Cervix;

RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP SPECIES=Mouse;  
RC SPECIES=Mouse;  
RA zeng Q., Tan Y.H., Hong W.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SPECIES FROM N.A.  
RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Stomach;  
RX MEDLINE=2108566; PubMed=1121785;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,  
RA Sakai L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -!- FUNCTION: May be involved in vesicular trafficking and  
CC neurotransmitter release  
CC  
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF106681; AAD43034.1; -  
DR EMBL; AF297660; AAG13413.1; -  
DR EMBL; AF297660; AAG13413.1; -  
DR EMBL; AF423223; BAB14474.1; -  
DR EMBL; AF498945; AAM21093.1; -  
DR EMBL; BC000896; AAH00896.1; -  
DR EMBL; AF035646; AAC29313.1; -  
DR EMBL; AK008725; BAB25858.1; -  
DR HSP; P05713; 3RAB;  
DR Genew; HGNC:9759; RAB10.  
DR MGD; MGI:105066; RAB10.  
DR InterPro; IPR003579; GTPase\_Rab.

DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR InterPro; IPR005225; Small\_Grp.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 199 199 S-geranylgeranyl cysteine  
FT (By similarity).  
FT LIPID 200 200 S-geranylgeranyl cysteine  
FT (By similarity).  
FT CONFLICT 106 106 N -> H (IN REF. 7).  
SQ SEQUENCE 200 AA; 22541 MW; 7F02B8E8E46E1E8 CRC64;  
Query Match 48.0%; Score 530.5; DB 1; Length 200;  
Best Local Similarity 55.8%; Pred. No. 2.2e-35;  
Matches 96; Conservative 29; Mismatches 29; Indels 1; Gaps 1;  
[6]  
Qy 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDYFQKTIIEVDGIKVI 59  
Db 1 MAKTYDLLFKLLIGDSGVGKTCVLFPSDDAFNTTIFSTIGDPKIKTVLQGGKIKL 60  
Qy 60 QIWDTAGQERYQTITKOYVRAQGIFLVYDISSRSYQHIMKWYSDVDVDEYAPGVOKILI 119  
Db 61 QIWDTAGQERPHITTSYRGMGIMLVYDITNGSKFENISKWRNIDEHANEDVERMLL 120  
Qy 120 GNKADEBQKRGVRCQQQLAKYGMDFYETSACTNLNLIKESFRLTELVLQ 171  
Db 121 GNKCDMDKRVKPGKGEQIAREHGIRFFETSANINIEKAFITLAEDILR 172  
RESULT 10  
R88B HUMAN  
ID R88B HUMAN STANDARD; PRT; 207 AA.  
AC Q92930; Q92933;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ras-related protein Rab-8B.  
GN RAB8B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RA Seki N., Saito T.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,





XX MEDLINE=91360267; PubMed=1886711;  
RA Nanno E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
RA Johnson K.J.;  
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
RT genes.";  
RL Oncogene 6:1347-1351(1991).  
RN [2]  
RN INTERACTION WITH MAP4K2.  
RP INRAIN=BALB/C; TISSUE=Melanoma;  
RC MEDLINE=96209873; PubMed=8643544;  
RA Ren M., Zeng J., De Lemos-Charandini C., Rosenfeld M., Adesnik M.,  
RA Sabatini D.D.;  
RT "In its active form, the GTP-binding protein rab8 interacts with a  
RT stress-activated protein kinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:5151-5155(1996).  
RN [3]  
RN INTERACTION WITH SYTL4.  
RP MEDLINE=22590467; PubMed=12590134;  
RA Fukuda M.;  
RT "Glp4-a/gramphilin-a inhibits dense-core vesicle exocytosis through  
RT interaction with the GTP-bound form of Rab27A in PC12 cells.";  
RL J. Biol. Chem. 278:15390-15396(2003).  
CC -1- FUNCTION: May be involved in vesicular trafficking and  
CC neurotransmitter release.  
CC -1- SUBUNIT: Interacts with MAP4K2 and SYTL4.  
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: S53270; AB19682.1; -  
DR HSPSP; P05713; 3RAB.  
DR MGD; MG1:96960; Mel.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRfams; TIGR00231; small GTP; 1.  
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;  
KW Proto-oncogene.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 203 203 S-geranylgeranyl cysteine  
FT (By similarity).  
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;  
Query Match 47.6%; Score 526; DB 1; Length 206;  
Best Local Similarity 56.4%; Pred. No. 5.2e-35;  
Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDGSVGKTCCLCRFTDNEFHSHSHITIGVDFPKMTIEVDGKVRQ 60  
Db 1 MAKYDYVLFKLLIGDGSVGKTCVLFPSDFAFNSTFTSTIGIDFKRTIELDGKRIKQ 60  
Qy 61 IWDTAGERYOTITKQYRRAGQIFLVYDISSERSYQIMKWSVDVEYAPGVQKILIG 120  
Db 61 IWDTAGQERFTITATYRGANGIMLVYDITNEKSFQIRNWRINRIBEHASADVVKMILG 120  
Qy 121 NKABEQRQVGRGQQLAKYEGMDVFTSACTNLIKESFTFL 165  
Db 121 NKDVNDKQVSKRGEKALDYGIKFMETSAKANINVENAFPTL 165

RESULT 13  
YPT2 SCHPO  
ID YPT2 SCHPO STANDARD; PRT; 200 AA.  
AC P17609;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ras-related protein ypt2.  
GN YPT2 OR SPAC9E9.07C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=90269232; PubMed=2112089;  
RA Hengst L., Lehmeier T., Gailwitz D.;  
RT "Structural and functional analysis of ypt2, an essential ras-related  
RT gene in the fission yeast Schizosaccharomyces pombe encoding a Sec4  
RT protein homologue.";  
RL EMBL J. 9:1957-1962(1990).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=90332438; PubMed=2115995;  
RA Fawell E., Hook S., Sweet D., Armstrong J.;  
RT "Novel YPT1-related genes from Schizosaccharomyces pombe.";  
RL Nucleic Acids Res. 18:4264-4264(1990).  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Skelton J., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprav B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Bartelli B.G., Nurse P.;  
RL "The genome sequence of Schizosaccharomyces pombe.";  
Nature 415:871-880(2002).  
CC -1- FUNCTION: Protein transport. Probably involved in vesicular  
CC traffic (By similarity).  
CC -1- MISCELLANEOUS: This protein is essential for cell viability.  
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.  
CC PROBABLE YEAST SEC4 HOMOLOG.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

```

DR EMBL; X52469; CAA36707.1; -.
DR EMBL; X52864; CAA37045.1; -.
DR EMBL; Z99262; CAB16405.1; -.
DR PIR; S12790; S12790.
DR HSP; P05713; 3RAB.
DR GeneDB SPombe; SPAC9E9.07c; -.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (PROBABLE).
FT LIPID 199 199 S-geranylgeranyl cysteine.
FT LIPID 200 200 (By similarity).
FT LIPID 200 200 S-geranylgeranyl cysteine.
FT LIPID 200 200 (By similarity).
SQ SEQUENCE 200 AA; 22755 MW; 2C658D153A290C30 CRC64;

Query Match 46.2%; Score 510; DB 1; Length 200;
Best Local Similarity 51.9%; Pred. No. 9.5e-34;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KOYDVLPRLLIGDSGVGKTCLLCRFTDNEPHSHSTIGVDPRKMTIEVDGIKVRIOIW 62
DB 4 KSYDYLIKLLIGDSGVGKSCLLRFSDFSFTPTFTTIGIDFKIRTIELDKRIKLIQIW 63

QY 63 DTAGQRYVOTTKQYVRAQGIPLVYDISSERSYOHIMKWSVDVEYAPGVQKILIGNK 122
DB 64 DTAGQRFRTITAYRGAMGILLVYDVTDKSPDNVRTWFSNVEQHASENVYKILIGNK 123

QY 123 ADEKQKQVGEQOQQLAKGYGMDFYTSACTNLNIKESFRLTELVLQAHKKELEGLRM 182
DB 124 CDCEQKQVFEQOQALADELGVKFLASAKTNVNDVDEAFFLA-----REIKKQKI 175

QY 183 RASNELA 189
DB 176 DAENEFS 182

RESULT 14
RAB1_BETVU
ID RAB1_BETVU STANDARD; PRT; 215 AA.
AC Q39433;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein RAB1B.
GN RAB1B.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
OC Caryophyllales; Amaranthaceae; Beta.
OX NCBI_TaxID=161934;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. D100 KS 38080;
RX MEDLINE=96307523; PubMed=8680960;
RA Dallery E., Quief S., Ben Jilany K.E., Kerckaert J.-P., Hagege D.;
RT "Molecular cloning and structural analysis of cDNAs that encode 3
small GTP-binding proteins from sugar beet.";
C. R. Acad. Sci., III, Sci. Vie 319:91-97(1996).
CC -/- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; Z49152; CAA89021.1; -.
DR PIR; T14565; T14565.
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT LIPID 212 212 S-geranylgeranyl cysteine.
FT LIPID 213 213 (By similarity).
FT LIPID 213 213 S-geranylgeranyl cysteine.
FT LIPID 213 213 (By similarity).
SQ SEQUENCE 215 AA; 23787 MW; 781CC5A17F29B8AA CRC64;

Query Match 46.1%; Score 509; DB 1; Length 215;
Best Local Similarity 47.8%; Pred. No. 1.2e-33;
Matches 100; Conservative 42; Mismatches 59; Indels 8; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEPHSHSTIGVDPRKMTIEVDGIKVRIOIWDT 64
DB 12 YDYLKLLIGDSGVGKSCLLRFSDFSFTPTFTTIGIDFKIRTIELDKRIKLIQIWDT 71

QY 65 AGQERYVOTTKQYVRAQGIPLVYDISSERSYOHIMKWSVDVEYAPGVQKILIGNKAD 124
DB 72 AGQERFRTITAYRGAMGILLVYDVTDESFNNRNWIRNIHQHASDNDVKNKILVGNKAD 131

QY 125 -BEQKQVGEQOQQLAKGYGMDFYTSACTNLNIKESFRLTELVLQAHKKELEGLRM 183
DB 132 MDESKAVPTAKQALADEYGIKFFETSAKTNLNVVEFFSIARDIKQ--RLADSDTROE 189

QY 184 ASNELALAELEEEBEGKEGPEGPANSSKTCWC 212
DB 190 AQSITIKPADQ-----SGNQAAAKSACC 213

RESULT 15
ARA3_ARATH
ID ARA3_ARATH STANDARD; PRT; 216 AA.
AC P28186;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein ARA-3.
GN ARA-3 OR AT3G46060 OR F12M12 30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta, cv. Columbia, cv. En-1, cv. Lapalmar,
and cv. Estland;
RX TISSUE=Leaf;
RX MEDLINE=92084144; PubMed=1749311;
RA Anai T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
Matsui M.;
RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins
of Arabidopsis thaliana.";
Gene 108:259-264(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

```

RA MEDLINE=21016720; PubMed=11130713;  
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,  
RA Fatmamb B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,  
RA De Senne V., Boutry M., Griwell L.A., Mache R., Puigdomenech P.,  
RA De Simone N., Choisine N., Artiguenave F., Robert C., Brottier P.,  
RA Wincker P., Catellico L., Weisenbach J., Saurin W., Quetier P.,  
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
RA Wurmbach E., Dronek H., Erle H., Jordan N., Bangert S.,  
RA Wedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,  
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,  
RA Reichelt J., Scharfe M., Schoen O., Baryes M., Terol J., Climent J.,  
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
RA Cooke R., Laudie M., Berger-Llauro C., Furnelle B., Masuy D.,  
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.-L., Tallon L.J., Jenkins J.,  
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
RA Pai G., Millscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakayama S., Nakazaki N., Shimp S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,  
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
RT thaliana";  
RL Nature 408:820-822(2000).  
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; D01025; BAA00830.1; --  
CC EMBL; ALJ35775; CAB90933.1; --  
CC PIR; JS0640; JS0640.  
CC HSSP; P05713; 3RAB.  
DR InterPro; IPR003579; GTPase Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Multigene family.  
FT NP\_BIND 22 29 GTP (BY SIMILARITY).  
FT NP\_BIND 70 74 GTP (BY SIMILARITY).  
FT NP\_BIND 128 131 GTP (BY SIMILARITY).  
FT DOMAIN 44 52 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 213 213 S-geranylgeranyl cysteine  
FT (By similarity).  
FT LIPID 214 214 S-geranylgeranyl cysteine  
FT (By similarity).  
SQ SEQUENCE 216 AA; 23835 MW; 09E9C19A7A44E705 CRC64;  
Query Match 46.0%; Score 508; DB 1; Length 216;  
Best Local Similarity 46.9%; Pred. No. 1.5e-33;  
Matches 100; Conservative 41; Mismatches 54; Indels 18; Gaps 3;  
QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDKMTIEVDGKIKVRIQIWDI 64  
DB 12 YDYLKULLIGDSGVGKSCLLRFSDGSFTTSFTTIGDFKIRTIELDKRIKLQIWDI 71  
QY 65 AGQRYOTITKQYVRAQGIPLVYVDISERSYQHMKWSDVDEYAPGEGVQKILGNKAD 124  
DB 72 AGQRFRTITTYRGGAMGILLVYDVTDESSFNIRNWRNQHSADNVNKKILVGNKAD 131

QY 125 -EEKQVQREGGQQLAKRYGMDFYETSACTNLNIKESFTRLTLVLQAHRKELEGLMR 183  
DB 132 MDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVVEVFPISIG-----RDIKQR 180  
QY 184 ASNELALAE-----LEEERKPEGPANSSKTC 210  
DB 181 LSDTDSRAEPATIKISQTDQAAGAGQATQKSAC 213

Search completed: March 15, 2004, 11:15:20  
Job time : 18 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:12:33 ; Search time 46 Seconds  
(without alignments)  
1454.127 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	72.2	168	11 Q91YW0	Q91yw0 mus musculus
2	573	51.9	211	4 Q81W89	Q81w89 homo sapien
3	563.5	51.0	200	13 Q7T3A4	Q7t3a4 brachydanio
4	557	50.4	211	5 Q9TYS2	Q9tys2 caenorhabdi
5	546.5	49.5	204	5 O15971	O15971 drosophila
6	541	49.0	207	11 Q8VCF6	Q8vcf6 mus musculus
7	528.5	47.8	203	11 Q8K3X5	Q8k3x5 rattus norv
8	528.5	47.8	207	5 O18338	O18338 drosophila
9	526.5	47.6	200	4 Q9H0T3	Q9h0t3 homo sapien
10	526	47.6	202	11 Q9DD03	Q9dd03 mus musculus
11	521	47.1	206	3 Q9HET4	Q9het4 aspergillus
12	519	47.0	216	10 Q24466	Q24466 arabidopsis
13	518	46.9	214	10 Q40218	Q40218 lotus japon
14	518	46.9	215	10 Q7XHP7	Q7xhp7 oryza sativ
15	515.5	46.7	216	10 Q40215	Q40215 lotus japon
16	515	46.6	216	10 Q9FUF1	Q9fuf1 arabidopsis

17	514	46.5	201	5 Q94148	Q94148 caenorhabdi
18	513.5	46.5	215	10 Q40219	Q40219 lotus japon
19	513	46.4	216	10 Q8WVF9	Q8wvf9 nicotiana t
20	512.5	46.4	214	10 Q84TA8	Q84ta8 oryza sativ
21	512	46.3	216	10 Q40177	Q40177 lycopersico
22	512	46.3	216	10 Q9LZD4	Q9lzd4 arabidopsis
23	511.5	46.3	212	10 Q40217	Q40217 lotus japon
24	510.5	46.2	216	10 Q8W3J4	Q8w3j4 nicotiana t
25	510	46.2	218	10 Q9SF91	Q9sf91 arabidopsis
26	508.5	46.0	215	10 Q41023	Q41023 pisum sativ
27	508	46.0	216	10 Q49844	Q49844 daucus caro
28	508	46.0	216	10 Q8W3J3	Q8w3j3 nicotiana t
29	507.5	45.9	222	10 Q96362	Q96362 brassica ca
30	507	45.9	202	3 Q96VL3	Q96vl3 colletotric
31	507	45.9	216	10 Q8L8Q4	Q8l8q4 arabidopsis
32	506.5	45.8	216	10 Q9SWV8	Q9swv8 lycopersico
33	506	45.8	216	10 Q41024	Q41024 pisum sativ
34	505	45.7	198	4 Q96GU4	Q96gu4 homo sapien
35	505	45.7	216	10 Q8W3J2	Q8w3j2 nicotiana t
36	502.5	45.5	215	10 Q41022	Q41022 pisum sativ
37	495.5	44.8	204	3 Q9C1Z5	Q9clz5 pichia past
38	492.5	44.6	215	10 Q41061	Q41061 pisum sativ
39	484	43.8	205	5 Q9BLF3	Q9blf3 entamoeba h
40	482	43.6	203	10 Q40569	Q40569 nicotiana t
41	482	43.6	205	5 Q26554	Q26554 schistosoma
42	481	43.5	203	10 Q8RU63	Q8ru63 oryza sativ
43	478	43.3	202	10 Q9ZRH6	Q9zrh6 petunia hyb
44	477	43.2	202	3 Q9HDT5	Q9hdt5 trichoderma
45	477	43.2	202	10 Q39845	Q39845 glycine max

## ALIGNMENTS

### RESULT 1

Q91YW0 ID Q91YW0 PRELIMINARY; PRT; 168 AA.  
AC Q91YW0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN 2310012G06RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strauberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
DR EMBL; BC013790; A013790.1; -.  
DR HSSP; P05713; 3RAH.  
DR MGD; MGI:1916865; 2310012G06RIK.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
DR GO; GO:0015031; P:protein transport; IEA.  
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
DR InterPro; IPR003579; GTPase Rab.  
DR InterPro; IPR001806; Ras trnafrmg.  
DR InterPro; IPR002078; Sig54\_interact.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRfams; TIGR00231; small\_GTP\_1.  
DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
KW Hypothetical protein; GTP-binding; Lipoprotein.  
SQ SEQUENCE 168 AA; 19399 MW; 4E2C2FB1C56BCDF8 CRC64;

Query Match	72.2%;	Score 798;	DB 11;	Length 168;
Best Local Similarity	99.3%;	Pred. No. 5.2e-60;		
Matches 152;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPKMTIEVDGKVKRIQ	60
Dd	1	MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPKMTIDVDGKVKRIQ	60
Qy	61	IWDTAGQERYQTITKQYRRAGQIFLVVDISSESYQHIMKMWSDVDEYAPGVQKILIG	120
Dd	61	IWDTAGQERYQTITKQYRRAGQIFLVVDISSESYQHIMKMWSDVDEYAPGVQKILIG	120
Qy	121	NKADEEQKRVGREQGOQLAKEYGMDFYETSAC	153
Dd	121	NKADEEQKRVGREQGOQLAKEYGMDFYETSAC	153

RESULT 2	
Q8IW89	PRELIMINARY; PRT; 211 AA.
AC	Q8IW89;
DT	01-MAR-2003 (TRENBLrel. 23, Created)
DT	01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE	01-OCT-2003 (TRENBLrel. 25, Last annotation update)
OS	Hypothetical protein (Fragment).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RA	Strausberg R.;
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; BC040679; AAH40679.1; --
DR	GO; GO:0003677; F:DNA binding; IEA.
DR	GO; GO:0005525; F:GTP binding; IEA.
DR	GO; GO:0003928; F:RAS small monomeric GTPase activity; IEA.
DR	GO; GO:0003930; F:RAS small monomeric GTPase activity; IEA.
DR	GO; GO:0003931; F:Rho small monomeric GTPase activity; IEA.
DR	GO; GO:0006886; P:intracellular protein transport; IEA.
DR	GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR	GO; GO:0000160; P:two-component signal transduction system [p. . .]; IEA.
DR	InterPro; IPR003579; GTPase Rab.
DR	InterPro; IPR003577; GTPase Ras.
DR	InterPro; IPR003578; GTPase_Rho.
DR	InterPro; IPR002041; RAN.
DR	InterPro; IPR001806; Ras trnsfrmg.
DR	InterPro; IPR002078; Sig54 interact.
DR	InterPro; IPR005225; Small_GTP.
DR	Pfam; PF00071; ras; 1.
DR	PRINTS; PR00449; RASTRNSFRMG.
DR	SMART; SM00175; RAB; 1.
DR	SMART; SM00176; RAN; 1.
DR	SMART; SM00173; RAS; 1.
DR	SMART; SM00174; RHO; 1.
DR	TIGRFAMs; TIGR00231; small GTP; 1.
DR	PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW	Hypothetical protein.
FT	NON TER 1
SQ	SEQUENCE 211 AA; 3BFA7CF89D4DF54F CRC64;

Query Match	51.9%;	Score 573;	DB 4;	Length 211;
Best Local Similarity	62.8%;	Pred. No. 9e-41;		
Matches 125;	Conservative 15;	Mismatches 39;	Indels 20;	Gaps 4

Qy	1	MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPKMTIEVDGKVKRIQ	60
Dd	4	MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSSHSHTIGVDFPKMTIEVDGKVKRIQ	63
Qy	61	IWDTAGQERYQTITKQYRRAGQIFLVVDISSESYQHIMKMWSDVDEYAPGVQKILIG	120
Dd	64	IWDTAGQERYQTITKQYRRAGQIFLVVDISSESYQHIMKMWSDVDEYAPGVQKILIG	113



RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; AF003572; AAF50924.1; -;  
 DR EMBL; AB061899; BAA21744.1; -;  
 DR EMBL; AY060425; AAL25464.1; -;  
 DR HSP; P05713; 3RAB.  
 DR FlyBase; FBgn0015789; Rab10.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0005252; F:GTP binding; IEA.  
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
 DR GO; GO:0015031; P:protein transport; IEA.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . . ; IEA.  
 DR InterPro; IPR003572; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR002078; Sig54 interact.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 204 AA; 23336 MW; 4E050761C6854920 CRC64;

Query Match 49.58; Score 546.5; DB 5; Length 204;  
 Best Local Similarity 49.88; Pred. No. 1.5e-38;  
 Matches 108; Conservative 47; Mismatches 50; Indels 9; Gaps 3;  
 QY 1 MAKOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGVDFFKMTIEVDGKIVRI 59  
 DB 1 MAKTYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGVDFFKMTIEVDGKIVRI 60  
 QY 60 QIMTAGQERYOTITTKOYRRAQIGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILI 119  
 DB 61 QIMTAGQERYOTITTKOYRRAQIGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILI 120  
 QY 120 GNKADEQKRVGREGQOQLAKYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEG 179  
 DB 121 GNKCDMTDKRVNKGEGEATAREHIGFIMETSAKSNINIERAFCELAEBAILD----KTSG 176  
 QY 180 LRMRASNELALAELEEEEGEPGANSKTC 210  
 DB 177 RESAENQERVILDRNQEAP----GYSKCC 203

RESULT 6  
 Q8VCF6 PRELIMINARY; PRT; 207 AA.  
 AC Q8VCF6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mel transforming oncogene (Derived from cell line NK14)-RAB8 homolog (Cell line NK14 derived transforming oncogene).  
 GN MEL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Liver;  
 RC Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Body, Cerebellum, and Retina;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 .RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; BC019990; AAH19990.1; -;  
 DR EMBL; AK076048; BAC36146.1; -;  
 DR EMBL; AK079306; BAC37603.1; -;  
 DR EMBL; AK080740; BAC38003.1; -;  
 DR HSP; P05713; 3RAB.  
 DR MGD; MGI:96960; Mel.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR002078; Sig54 interact.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 207 AA; 23668 MW; AC89DC85588FB8F8 CRC64;

Query Match 49.08; Score 541; DB 11; Length 207;  
 Best Local Similarity 50.38; Pred. No. 4.6e-38;  
 Matches 99; Conservative 49; Mismatches 49; Indels 0; Gaps 0;  
 QY 1 MAKOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGVDFFKMTIEVDGKIVRI 60  
 DB 1 MAKTYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGVDFFKMTIEVDGKIVRI 60  
 QY 61 IWTAGQERYOTITTKOYRRAQIGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILI 120  
 DB 61 IWTAGQERYOTITTKOYRRAQIGIFLVYDISSERSYOHIMKWSVDVDEYAPGVQKILI 120  
 QY 121 NKADEQKRVGREGQOQLAKYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEG 180  
 DB 121 NKCDVNDKRVSKERGKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKEG 180  
 QY 181 RMRASNELALAELEEE 197  
 DB 181 SPQSSHGKVTVEQQK 197

RESULT 7  
 Q8KX35 PRELIMINARY; PRT; 203 AA.  
 AC Q8KX35;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE GTP-binding protein RAB13.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RC Lau A.S.N., Cheng C.Y., Mruk D.D.;  
 RT "Rab13 participates in Sertoli cell tight junction assembly in the testis."  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF525280; AA082588.1; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0005252; F:GTP binding; IEA.  
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
 DR GO; GO:0003930; F:RAB small monomeric GTPase activity; IEA.  
 DR GO; GO:0003931; F:Rho small monomeric GTPase activity; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . . ; IEA.  
 DR InterPro; IPR006688; ARF.



```
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR003577; GTPase Ras.
DR InterPro; IPR003578; GTPase rho.
DR InterPro; IPR002041; RAN.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00177; ARE; 1.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAN; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
DR GTP-binding.
SQ SEQUENCE 203 AA; 22901 MW; B6687462C6EA95A4 CRC64;
Query Match 47.8%; Score 528.5; DB 11; Length 203;
Best Local Similarity 48.1%; Pred. No. 5.2e-37;
Matches 100; Conservative 53; Mismatches 46; Indels 9; Gaps 2;
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKVRIQ 60
Db 1 MAKAYDLFLKLLIGDSGVGKTCILIRFAEDNFNSTVISTIGIDFKIRTVEIEGKRIKQ 60
Qy 61 IWDTAGERYQTITKQYRRAGQIFLYVDISSERTSYOHIMKWSVDVEYAPGKVLIG 120
Db 61 VMDTAGQERFKTITTYIRGANGILVYDITDEKSFENIQNMKSIIKENASAGVERLLG 120
Qy 121 NKADEOKRGVREGOQOLAKYGMDFVTSACTNLNLIKESFRTLVLQAHKKELEGL 180
Db 121 NKCDMEARKVQORQAEHLAREHRIREFTSKSVNVDERAFSSLDIL----LKTGR 176
Qy 181 RMRASNELALAELEBEKGPENANSK 208
Db 177 RSGNSKSPSSDILKVSDDK-----NSNK 199
RESULT 8
O18338 PRELIMINARY; PRT; 207 AA.
AC O18338
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAB8 protein (Lb44762p).
GN RAB8 OR CG8287.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bereman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo J.B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
```

```
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=OREGON R; TISSUE=Head;
RX MEDLINE=97228579; PubMed=9074639;
RA Satoh A.K., Tokunaga F., Ozaki K.;
RT "Rab proteins of Drosophila melanogaster: novel members of the Rab-
protein family.";
RL FBS Lett. 404:65-69(1997).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AE003515; AAF49101.1; -.
DR EMBL; D84347; BAA21711.1; -.
DR EMBL; AY069671; AAL39816.1; -.
DR HSSP; P05713; 3RAB
DR FlyBase; FBgn0015796; RAB8
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 207 AA; 23691 MW; BE9DE812C77DCF09 CRC64;
Query Match 47.8%; Score 528.5; DB 5; Length 207;
Best Local Similarity 53.8%; Pred. No. 5.3e-37;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKVRIQ 60
Db 1 MAKTYDVLFRLLIGDSGVGKTCILIRFSDAFNTFTISTIGIDFKIRTDLNKKIKQ 60
Qy 61 IWDTAGERYQTITKQYRRAGQIFLYVDISSERTSYOHIMKWSVDVEYAPGKVLIG 120
```





```
Db 12 YDYLKLLLGDSGVGKSCLLRFSDSFTTSTFTTIGDKIRTIELDKGRKIKLQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHSDSNVKNILVGNKAD 131
Qy 125 -BEQKRVGRGQGLAKKEYGMDYFETSACTNINIKESFTRLTLVLQA-----HRKELE 178
Db 132 MDESKRAVPTSKQALADEYGMKFETSAKTNLNVVEVFFSIADIKQRLADTDARAEPQ 191
Qy 179 GLMRASNE 187
Db 192 TIKINQSDQ 200

RESULT 13
Q40218 ID Q40218 PRELIMINARY; PRT; 214 AA.
AC Q40218;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAB8D.
GN RAB8D.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -I- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73947; CAA98175.1; -.
DR HSP; P05713; 3RAB.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005252; F:GTP binding; IEA.
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 214 AA; 23627 MW; 5E1AGE83505E50D7 CRC64;

Query Match 46.9%; Score 518; DB 10; Length 214;
Best Local Similarity 48.6%; Pred. No. 4.3e-36;
Matches 103; Conservative 38; Mismatches 55; Indels 16; Gaps 3;

Qy 5 YDVLFRLLLGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKIKVRIQIWD 64
Db 12 YDYLKLLLGDSGVGKSCLLRFSDSFTTSTFTTIGDKIRTIELDKGRKIKLQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHSDSNVKNILVGNKAD 131
Qy 125 -BEQKRVGRGQGLAKKEYGMDYFETSACTNINIKESFTRLTLVLQA-----HRKELE 178
Db 132 MDESKRAVPTSKQALADEYGMKFETSAKTNLNVVEVFFSIADIKQRLADTDARAEPQ 191
Qy 179 GLMRASNE 187
Db 192 TIKINQSDQ 200

RESULT 14
Q7XHP7 ID Q7XHP7 PRELIMINARY; PRT; 215 AA.
AC Q7XHP7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0565A07.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005194; BAC80082.1; -.
SQ SEQUENCE 215 AA; 23993 MW; F3DAD303C533C7EE CRC64;

Query Match 46.9%; Score 518; DB 10; Length 215;
Best Local Similarity 47.5%; Pred. No. 4.4e-36;
Matches 103; Conservative 42; Mismatches 48; Indels 24; Gaps 4;

Qy 5 YDVLFRLLLGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKIKVRIQIWD 64
Db 12 YDYLKLLLGDSGVGKSCLLRFSDSFTTSTFTTIGDKIRTIELDKGRKIKLQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHSDSNVKNILVGNKAD 131
Qy 125 -BEQKRVGRGQGLAKKEYGMDYFETSACTNINIKESFTRLTLVLQAHRKELEGLMR 183
Db 132 MDESKRAVPTSKQALADEYGMKFETSAKTNLNVVEVFFSIADIKQ-----179
Qy 184 ASNELALAELEEE-----GKPEGPANSS---KTCWC 212
Db 180 ---RLAETDSKTDRTIKINKEGDAEATTLQKSACC 213

RESULT 15
Q40215 ID Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAB8A.
GN RAB8A.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
```

```
Db 132 MDESKRAVPTSKQALADEYGMKFETSAKTNLNVVEVFFSIADIKQRLADTDHKAAPT 191
Qy 179 GLMRASNELALAELEEECKPEGPANSSKTC 210
Db 192 TLKINQDSAGAGE-----AANKSSCC 213

RESULT 14
Q7XHP7 ID Q7XHP7 PRELIMINARY; PRT; 215 AA.
AC Q7XHP7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0565A07.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005194; BAC80082.1; -.
SQ SEQUENCE 215 AA; 23993 MW; F3DAD303C533C7EE CRC64;

Query Match 46.9%; Score 518; DB 10; Length 215;
Best Local Similarity 47.5%; Pred. No. 4.4e-36;
Matches 103; Conservative 42; Mismatches 48; Indels 24; Gaps 4;

Qy 5 YDVLFRLLLGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKIKVRIQIWD 64
Db 12 YDYLKLLLGDSGVGKSCLLRFSDSFTTSTFTTIGDKIRTIELDKGRKIKLQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHSDSNVKNILVGNKAD 131
Qy 125 -BEQKRVGRGQGLAKKEYGMDYFETSACTNINIKESFTRLTLVLQAHRKELEGLMR 183
Db 132 MDESKRAVPTSKQALADEYGMKFETSAKTNLNVVEVFFSIADIKQ-----179
Qy 184 ASNELALAELEEE-----GKPEGPANSS---KTCWC 212
Db 180 ---RLAETDSKTDRTIKINKEGDAEATTLQKSACC 213

RESULT 15
Q40215 ID Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAB8A.
GN RAB8A.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:58:09 ; Search time 2486 Seconds  
(without alignments)  
2546.574 Million cell updates/sec

Title: US-09-817-198c-2  
Perfect score: 1105  
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPto\_spool\_p/US09817198/runat\_15032004\_101746\_20571/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastcap -SUFFIX=rat -MINMATCH=0.1 -LOOPFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CIGN 1.1 4237 @runat\_15032004\_101746\_20571 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	996	90.1	1091	10	BF160330	BF160330 601768601
2	988	89.4	1004	13	BUS17848	BUS17848 AGENCOURT
3	953.5	86.3	932	10	BF535642	BF535642 602054039
4	922	83.4	616	12	BI648588	BI648588 603277781
5	915.5	82.9	759	13	BU343508	BU343508 603523088
6	914	82.7	992	10	BF101730	BF101730 601753464
7	910	82.4	902	10	BF178163	BF178163 601809083
8	885	80.1	699	12	BI557933	BI557933 603237549
9	884	80.0	515	29	AY420960	AY420960 Homo sapi
10	882	79.8	627	9	AL863254	AL863254 AL863254
11	880.5	79.7	773	14	CK030691	CK030691 AGENCOURT
12	876	79.3	515	29	AY420961	AY420961 Pan trogl
13	872	78.9	515	29	AY420962	AY420962 Mus muscu
14	858.5	77.7	796	10	BF966292	BF966292 602286692
15	846	76.6	623	14	CB578521	CB578521 AMGNNUC:N
16	843	76.3	882	13	BU841314	BU841314 AGENCOURT
17	838	75.8	1100	10	BF181167	BF181167 601805830
18	833.5	75.4	3151	11	CNSLT118T	EX248046 human ful
19	829	75.0	616	12	BG247902	BG247902 602359819
20	816	73.8	1121	10	BF182001	BF182001 601805936
21	815	73.8	481	13	EX283729	EX283729 EX283729
22	799	72.3	814	12	BG762967	BG762967 602735022
23	792	71.7	660	12	BG974411	BG974411 602844184
24	777	70.3	911	12	BI649317	BI649317 603278051
25	756	68.4	618	10	BF150485	BF150485 uy86D05.Y
26	733	66.3	1767	12	BM462147	BM462147 AGENCOURT
27	727	65.8	485	10	BE913096	BE913096 601668580
28	706	63.9	667	12	BI685808	BI685808 603309467
29	704	63.7	843	12	BG865609	BG865609 602783763
30	693.5	62.8	577	12	BG703353	BG703353 602685115
31	687	62.2	652	9	AL886867	AL886867 AL886867
32	686	62.1	465	10	BF565114	BF565114 UI-R-B01-
33	670	60.6	479	10	BF320679	BF320679 uz55d10.Y
34	666	60.3	543	10	BG019552	BG019552 daa21e04.Y
35	663	60.0	539	10	BG022618	BG022618 uy51f07.Y
36	658.5	59.6	693	10	BE287683	BE287683 601093322
37	653	59.1	468	13	BY276887	BY276887 BY276887
38	651	58.9	485	13	BY255176	BY255176 BY255176
39	651	58.9	823	12	BM017178	BM017178 603643780
40	650	58.8	479	13	BY256967	BY256967 BY256967
41	649	58.7	385	12	BM936454	BM936454 UI-M-BH2
42	645	58.4	487	13	BY247617	BY247617 BY247617
43	639	57.8	416	13	BY269783	BY269783 BY269783
44	639	57.8	468	13	BY260718	BY260718 BY260718
45	638	57.7	483	13	BY243237	BY243237 BY243237

ALIGNMENTS

RESULT 1  
BF160330  
LOCUS BF160330  
DEFINITION 601768601F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3987745 5',  
mRNA sequence.  
ACCESSION BF160330  
VERSION BF160330.1 GI:11040541  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1091)

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM9195 row: b column: 02  
 High quality sequence stop: 654.  
**FEATURES** Location/Qualifiers  
 source 1..1091  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3987745"  
 /tissue\_type="spontaneous tumor, metastatic to mammary."  
 Stem cell origin.  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Lu29"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

**Alignment Scores:**  
 Pred. No.: 4.85e-108 Length: 1091  
 Score: 996.00 Matches: 198  
 Percent Similarity: 94.42% Conservative: 5  
 Best Local Similarity: 92.09% Mismatches: 9  
 Query Match: 90.14% Indels: 3  
 DB: 10 Gaps: 0

US-09-817-198C-2 (1-212) x BF160330 (1-1091)

**Qy** 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20  
**Db** 76 ATGCGGAAACAGTACGATGTCTCTCCGGCTTCTGCTGATCGGGACTCGGGGTTGGC 135  
**Qy** 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40  
**Db** 136 AAGACATGCTGCTGTGCTCCGCTTCCCGACACAGAGTTCCACTCTCGCATATCTCCACC 195  
**Qy** 41 IleGlyValAspPheLys-MetLysThrIleGluValAspGlyIleLysValArgIleG1 60  
**Db** 196 ATCGGTGTTGACTTTACGATGAGACTATCGAGTAGAGCGCATCAAGTAGAGATACA 255  
**Qy** 60 nileTtrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAr 80  
**Db** 256 GATTGGGACACAGCAGCGGAGGAGAGGTACACACTATCACAAAGCAGTACTATCGCGC 315  
**Qy** 80 gAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMe 100  
**Db** 316 AGCCAGGGGAATATTTTAGTCTACGACATTAGCAGTGAGCGCTCTATCAGCATATCAT 375  
**Qy** 100 tLys-TTpValSerAspValAspGlnTyrAlaProGluGlyValGlnLysIleLeuIleG 120  
**Db** 376 GAACGTGGTTCAGTGACGTGGATGAGTACGCTCCAGAGGAGTCCAGAGATCTTAATG 435  
**Qy** 120 lyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuA 140  
**Db** 436 GGAATAAGGCTGATGACGAGCAGAAACCGCAGGTGGGAGAGAGAGAGCGGCGAGCAGCTG 495  
**Qy** 140 lalysGluTyrGlyMetAspPheTyrGluThrSerAlaCyeThrAsnLeuAsnIleLysG 160  
**Db** 496 CTAAGGATGCGGATGGACTTCTACGAAACAGTGGCTGCGCCACCACTTAATATTAAAG 555

**Qy** 160 luSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyL 180  
**Db** 556 AGTCCTTCACTGCTGACGAGCTGCTGCTGAGGCCACAGGAAAGAGCTGGATGTC 615  
**Qy** 180 euArgMetArgAlaSer-AsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLys 199  
**Db** 616 TCCGAACACGTGCCAGCAACGAGCTGCGACTGGCGAGCTGCAGGAGGAGGAGGAGG 675  
**Qy** 200 ProGluGlyProAlaLeuSerSerLysThrCysTyrCys 212  
**Db** 576 CCTGGGGCCCGACCAATCTTCAAGGACTGGCTGGTGC 714

## RESULT 2

BU517848

LOCUS

DEFINITION

AGENCOURT 10165969 NCI CGAP Mam2 Mus musculus cDNA clone

IMAGE:6515476 5', mRNA sequence.

BU517848

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM14092 row: h column: 05

High quality sequence stop: 586.

## FEATURES

source

Location/Qualifiers

1..1004

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N-3"

/db\_xref="taxon:10090"

/clone="IMAGE:6515476"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="5 months"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Mam2"

/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

Alignment Scores:

Pred. No.: 3.85e-107 Length: 1004

Score: 988.00 Matches: 197

Percent Similarity: 98.06% Conservative: 5

Best Local Similarity: 95.63% Mismatches: 1

Query Match: 89.41% Indels: 3

DB: 13 Gaps: 0

US-09-817-198C-2 (1-212) x BU517848 (1-1004)

**Qy** 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20**Db** 153 ATGCGAAACAGTACGATGTCTTCCGGCTACTGCTGATCGGGACTCGGGGTTGGC 212**Qy** 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40**Db** 213 AAGACATGCTGCTGTGCTCCGCTTCCCGACACAGAGTTCCACTCTCGCATATCTCCACC 272



Qy 41 ilecylvalasphepelymetlyethrileuvalaspGlylelyValArgileGln 60  
 |||||  
 Db 273 ATCGGTGTCAGCTTTAAGATGAAGACTATCGATAGACGGCATCAAAAGTGAAGATACAG 332  
 |||||  
 Qy 61 IletpaspthrAlaGlyGlnGluArgTyGlnThrileThrlyleThrlySerGlyArg 80  
 |||||  
 Db 333 ATTTGGGACACAGCAGCGGCGAGAGATACAGACTATCAAAAGCAGTACTATCGCGCA 392  
 |||||  
 Qy 81 AlaGlnGlylePheLeuValTyAspIleSerSerGluArgSerTyGlnHisIleMet 100  
 |||||  
 Db 393 GCCCAGGGAATATTTAGTCTACGACATAGAGTGGGCTCCTATCAGCATATCATG 452  
 |||||  
 Qy 101 LysTrpValSerAspValaspGluTyAlaprogGlyValGlnLysIleLeuIleGly 120  
 |||||  
 Db 453 AAGTGGGTGAGTACGTTGGATGAGTACGCTCCAGAGGAGTCCAGAGAGATCTTAATTGGG 512  
 |||||  
 Qy 121 AsnlyslalaspGluGlnLysArgGlnValGlyArgGlnGlnGlnLeuAla 140  
 |||||  
 Db 513 AATAAGGCTGATGAAGAGCAGAAACCGCAGTGGGGAGAGCAGGCGCAGCTGGCT 572  
 |||||  
 Qy 141 LysGluTyGlyMetAspPheTyGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 |||||  
 Db 573 AAGAGTACGCGATGGACTTCTCGAAACAGTCCCTGCACCAACCTTAATATTAAAGAG 632  
 |||||  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 |||||  
 Db 633 TCCTTCACTCGTCTGACGGAGTGGTGTGCGAGGCCACAGGAAAGAGTGGATGGTCTC 692  
 |||||  
 Qy 181 ArgMetArgAlaSerAsnGluLeuAlaLeu-AlaGluLeuGluGlu-GluGluGly-Lys 199  
 |||||  
 Db 693 CGAACACGTGCCAGCAGCTCGCACTGGGCCAGCTGGAAGAGGACCCAGGCCAAA 752  
 |||||  
 Qy 200 ProGluGlyPro 203  
 |||||  
 Db 753 CTTGAGGGGGCCC 764  
 |||||

RESULT 3  
 BF535642 602054039F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4193456 5',  
 LOCUS mRNA sequence.  
 BF535642 932 bp mRNA linear EST 11-DEC-2000

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Mus musculus (house mouse)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapsb@email.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9525 row: m column: 09

High quality sequence stop: 685.

Location/Qualifiers

1. 932

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4193456"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_SG2"

FEATURES  
 source

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 616)

/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1:  
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
 dT. Average insert size 1.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."

# ORIGIN

## Alignment Scores:

Pred. No.: 4.57e-103 Length: 932  
 Score: 953.50 Matches: 187  
 Percent Similarity: 94.63% Conservative: 7  
 Best Local Similarity: 91.22% Mismatches: 9  
 Query Match: 86.29% Indels: 2  
 DB: 10 Gaps: 1

US-09-817-198C-2 (1-212) x BF535642 (1-932)

Qy 1 MetAlaLysGlnTyAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20  
 |||||  
 Db 145 ATGCGGAAACAGTACGATGTGCTTCCGGCTACTGCTGATCGGGGACTCGGGGTTGGC 204  
 |||||  
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40  
 |||||  
 Db 205 AAGACATGCTGTGTGCTGCTTCCGCAACAGAGTTCCACTCTCGCATATCTCCACC 264  
 |||||  
 Qy 41 IleGlyValaspPheLysMetLysThrIleGluValaspGlyIleLysValArgIleGln 60  
 |||||  
 Db 265 ATCGGTGTTGACTTTAAGATGAAGACTATCGATGATAGCGCATCAAAAGTGAGAATACAG 324  
 |||||  
 Qy 61 IletpaspThrAlaGlyGlnGluArgTyGlnThrileThrlyleThrlySerGlyArg 80  
 |||||  
 Db 325 ATTTGGGACACAGCAGGCGAGGAGGTACAGACTATCAAAAGCAGTACTATCGCATATCATG 444  
 |||||  
 Qy 81 AlaGlnGlylePheLeuValTyAspIleSerSerGluArgSerTyGlnHisIleMet 100  
 |||||  
 Db 385 GCCCAGGGAATATTTAGTCTACGACATTTAGCAGTGGCTCTCTATCAGCATATCATG 444  
 |||||  
 Qy 101 LysTrpValSerAspValaspGluTyAlaprogGlyValGlnLysIleLeuIleGly 120  
 |||||  
 Db 445 AAGTGGGTGAGTACGTTGGATGAGTACGCTCCAGAGGAGTCCAGAGAGTCTTAATTGGG 504  
 |||||  
 Qy 121 AsnlyslalaspGluGlnLysArgGlnValGlyArgGlnGlnGlnLeuAla 140  
 |||||  
 Db 505 AATAAGGCTGATGAAGAGCAGAAACCGCAGTGGGGAGAGCAGGCGCAGCTGGCT 564  
 |||||  
 Qy 141 LysGluTyGlyMetAspPheTyGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 |||||  
 Db 565 AAGAGTACGCGATGGACTTCTACGAAACAGTGGCTCGCACAACCTTAATATTCAAGAG 624  
 |||||  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 |||||  
 Db 625 TCCTTCACTCGTCTGACGGAGTGGTGTGCGAGGCCACAGGAAAGAGTGGATGGGCTC 684  
 |||||  
 Qy 181 ArgMetArgAla-SerAsnGluLeuAlaLeuAla---GluLeuGluGluGluGlyLy 199  
 |||||  
 Db 685 CGAACACGTGCCCGCAACAGCTCGCACTGGGCGAGACTGTGAGAGCAGCAGCAAGAGCGA 744  
 |||||  
 Qy 199 sProGluGlyPro 203  
 |||||  
 Db 745 ACCTGATGGGCCA 757  
 |||||

## RESULT 4

BF535642

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 616)

BI648588 616 bp mRNA linear EST 12-SEP-2001  
 60327781F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5317917 5',  
 mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 616)

# **AUTHORS** **TITLE** **JOURNAL** **COMMENT**

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M1806 row: a column: 22  
High quality sequence start: 2  
High quality sequence stop: 614.  
Location/Qualifiers

## **FEATURES**

source

1..616  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129,C57BL/6J,FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5317917"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam3"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

## **ORIGIN**

### **Alignment Scores:**

Pred. No.: 1.39e-99 Length: 616  
Score: 922.00 Matches: 180  
Percent Similarity: 98.92% Conservative: 3  
Best Local Similarity: 97.30% Mismatches: 2  
Query Match: 83.44% Indels: 1  
DB: 12 Gaps: 0

US-09-817-198c-2 (1-212) x B1648588 (1-616)

QY 28 PheThrAspGluPheHisSerSerHisLeSerThrIleGlyValAspPheLysMet 47  
DB 25 TTCACCGACACGAGTGTCCATCTCTCGCATATCTCCACCATCGGTGTGACTTTAAGATG 84  
QY 48 LysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGln 67  
DB 85 AAGACTATCGATGTAGACGGCATCANAGTCAGATACAGATTGGGACACAGCGGCAG 144  
QY 68 GluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuVal 87  
DB 145 GAGAGGTACAGACTATCACAAAGCAGTACTATCGCGAGCCCA-GGAATATTTTAGTC 203  
QY 88 TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAsp 107  
DB 204 TACGACATTAGCAGTGGCGCTCTATCAGCATATCATGAAGTGGGTCAGTCAGCTGGAT 263  
QY 108 GluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGln 127  
DB 264 GAGTACGCTCCAGAGAGGTCAGAGATCTTAATTTGGATAGGCTGATGAGAGCAG 323  
QY 128 LysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyrGlyMetAspPhe 147  
DB 324 AAACGGCAGGTGGGAGAGACGAGCGGCAGCGTGGCTAAGGAGTACGGCATGCTTC 383  
QY 148 TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167  
DB 384 TACGAACAAGTGCCTGCACCAACCTTAATTAAGAGAGTCCTTCACTCTGCTGACGAG 443  
QY 168 LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGlu 187

DB 444 CTGGTGTGTCAGGCCACAGAAAGAGCTGGATGGTCTCCGAACACGTGCCAGCAACGAG 503  
QY 188 LeuAlaIeuAlaGluLeuGluGluGluGluGlyProGluGlyProAlaAsnSerSer 207  
DB 504 CTCGCATGTCGCCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 563  
QY 208 LysThrCysTrpCys 212  
DB 564 AAGACTGTCTGGTGC 578

## **RESULT 5**

BU343508

LOCUS

DEFINITION

BU343508

VERSION

BU343508.1

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..759

source

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="Compton Line 151"

/db\_xref="taxon:9031"

/clone="CHEST465022"

/sex="female"

/tissue\_type="cerebellum"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="CSECHN68"

/note="Organ: brain; Vector: pBluescript II KS(+); Site\_1:

EcoRI; Site\_2: NotI; this normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

## **ORIGIN**

Alignment Scores:

Pred. No.: 1.16e-98 Length: 759

Score: 915.50 Matches: 176

Percent Similarity: 90.91% Conservative: 14

Best Local Similarity: 84.21% Mismatches: 18

Query Match: 82.85% Indels: 1



```

RESULT 7
BF178163          902 bp      mRNA      linear      EST 31-OCT-2000
LOCUS             601809083F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4039701 5',
DEFINITION        mRNA sequence.
ACCESSION         BF178163
VERSION           BF178163.1 GI:11056305
KEYWORDS          Mus musculus (house mouse)
SOURCE            Mus musculus
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 902)
AUTHORS          NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM9320 row: f column: 22
                  High quality sequence start: 2
                  High quality sequence stop: 603.
FEATURES          Location/Qualifiers
                   source
                     1..902
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4039701"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI CGAP Mam5"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"

ORIGIN
Alignment Scores:
Pred. No.:        6.91e-98      Length:      902
Score:            910.00        Matches:    177
Percent Similarity: 97.28%      Conservative: 2
Best Local Similarity: 96.20%   Mismatches: 4
Query Match:      82.35%       Indels:     1
DB:               10           Gaps:       0

US-09-817-198C-2 (1-212) x BF178163 (1-902)

Qy 30 AsphenGluPheHisSerSerHisIleSerThrIleGlyValAspPheLysMetLysThr 49
Db 14 GACAAACGAGTTCCACTCTCCGATATCTCCACCATCGGTGTGACTTTAAGATGAAGACT 73

Qy 50 IleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArg 69
Db 74 ATCGAGTAGACGGCATCAAGTGTATATACAGATTTGGGACACAGAGGGCAGAGAGG 133

Qy 70 TyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlnIlePheLeuValTyrAsp 89
Db 134 TACCAGACTATCAAAAGCAGTACTATCGGGAGGCCAGGGAATATTTTAGTCTACGAC 193

Qy 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyr 109
Db 194 ATTAGCAGTAGCGGCTCTCTATCAGATATCATGAAGTGGGTGAGTGGTGGTGGTGGT 253

Qy 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129

```

```

Db 254 GTCCACAGAGGAGTCCAGAGATCCTTAATTGGGAATAAGCTGATGAAGACGAGAACGG 313
Qy 130 GlnValGlyArgGlnGlnGlnGlnLeuAlaLys- GluTyrGlyMetAspPheTyrG1 149
Db 314 CAGTGGGAGAGAGAGAGGGGAGAGCTGGCTAAGCGAGTACGGCATGGCATCTTACGA 373
Qy 149 uThrSerAlaCyThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVa 169
Db 374 AACAGTGGCTGCACCAACCTTAATATAAAGAGTCCTTCACTCGTCTGACGAGCTGGT 433
Qy 169 lLeuGlnAlaHisArgLysGluLeuGlnGlyLeuArgMetArgAlaSerAsnGluLeuAl 189
Db 434 GCTGCGAGGCCACAGACAGAGCTGGTGTCTCCGAACACGTGCCAGCAACGAGCTCGC 493
Qy 189 aLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysTh 209
Db 494 ACTGGCGAGCTGGAGAGGAGGACGAGCAACCTGAGGGCCGACCAACTCTTCNAAGAC 553
Qy 209 rCysTrpCys 212
Db 554 CTGCTGGTGC 563

RESULT 8
BI557933
LOCUS             603237549F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5290284 5',
DEFINITION        mRNA sequence.
ACCESSION         BI557933
VERSION           BI557933.1 GI:15445247
KEYWORDS          Mus musculus (house mouse)
SOURCE            Mus musculus
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 699)
AUTHORS          NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM11734 row: b column: 13
                  High quality sequence stop: 698.
FEATURES          Location/Qualifiers
                   source
                     1..699
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129, C57BL/6J, FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5290284"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI CGAP Mam3"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."

ORIGIN
Alignment Scores:
Pred. No.:        4.5e-95      Length:      699
Score:            885.00        Matches:    172
Percent Similarity: 98.87%      Conservative: 3

```



Contact: Taylor R  
Sanger Institute  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS SEQUENCE ID: TEGg119f07.plkSP6  
Sequencing primer: SP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCSI07 with EcoRI at the  
5' end and NotI at the 3' end.  
Vector: pCSI07; Site 1: EcoRI; Site 2: NotI  
Host: *Escherichia coli* XL1-blue.

**FEATURES**  
**SOURCE**

```

1. 627
locus=700,quad=101
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TGg119f07"
/dev_host="egg"
/lab_host="Echerichia coli XL1-blue"
/clone_lib="XGC-egg"
/notes="Vector: pCS107; Site 1: EcoRI;
was oligo d primed from sug of poly
EcoRI-NotI cut cDNA was then ligated
EcoRI at the 5' end and NotI at the 3'
end."
8.56e-95      Length:      627
882.00      Matches:      167
96.69%      Conservative:      8
92.27%      Mismatches:      6
79.82%      Indels:      0
9      Gaps:      0

```

## ORIGIN

Alignment Scores:	
Pred. No.:	8,668-95
Score:	882.00
Percent Similarity:	96.69%
Best Local Similarity:	92.27%
Query Match:	79.82%
DB:	9
Gaps:	0
Indels:	0
Mismatches:	6
Conservative:	8
Matches:	167
Length:	627

US-09-817-198C-2 (1-212) x AL863254 (1-627)

Qy		1	MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuileGlyAapSerGlyValGly	20
Db		83	AyGGCTAGACAAATATGATGTCTTCGGGTGCTCTCATCGGGACTCCGGCGTCGG	142
Qy		21	LysThrCysLeuLeuCysArgPheThrAspAenGluPheHisSerHisIleSerThr	40
Db		143	AAGACTGTTTGCTCTGCAGGTTCAAGACAACGAATCCACCCTTCCCACATCTCCACT	202
Qy		41	IleGlyValAspPheLysMetIlysthrlleGluValAspGlyIleLysValArgIleGln	60
Db		203	ATAGGAGTTGACTTCAAATGAAGACAATAGAGGTTCATGGATCAAAGTAAGAATACAA	262
Qy		61	IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg	80
Db		263	ATCTCGGCACACAGCTGGTCAAGAACGATACACAGACAATTACCAACACAGTACTATAGAAGA	322
Qy		81	AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet	100
Db		323	GCAcAGGTATATTTCTTAGTCTATGACATACAGTGNACGCTCATACAGCATATTAATG	382
Qy		101	LysTrpValSerAspValAspGlnTyrAlaProGluGlyValGlnLysIleLeuileGly	120
Db		383	AAATGGCCAAAGTGATGTAGTAGTACGCCCGACCGAGGAGTTCAAAAAGATTTTAATTTGGG	442
Qy		121	AnnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla	140
Db		443	ANCAAAGCTGATGAAGAGCAAAAAACGACAGGTTGGAAAAATTCAGGGGATGCACTTGCA	502
Qy		141	LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAnnLeuAnnIleLysGlu	160
Db		503	GAGAGTATGGAAATGGACTTCTTTGAGACAAAGTGCCTTGCACTAATTACATATANAAGGAG	562
Qy		161	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180

563	TCATTCCTACGCTGCACAGAGCTGGTCTGCTGATGGCTCACAAAAGGGAGCTGGAGGGCTTG	622
Db		
Qy	181	Arg 181
Db	623	CGA 625
RESULT 11		
CK030691		
LOCUS		
DEFINITION	773 bp mRNA linear EST 26-NOV-2003	
	AGENCOURT.16648788 NIH MGC 203 Mus musculus cDNA clone	
	IMAGE:30787643 5', mRNA sequence.	
ACCESSION	CK030691	
VERSION	CK030691.1	GI:38556615
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
JOURNAL	1 (bases 1 to 773)	
COMMENT	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: <a href="mailto:cgabbs-femail.nih.gov">cgabbs-femail.nih.gov</a> Tissue Procurement: Naryan Bhat cDNA Library Preparation: Express Genomics cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: NDAMI100 row: 9 column: 12 High quality sequence stop: 706.	

## FEATURES

1. 773  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30787643"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 203"  
/note="Organ: placenta; Vector: pExpress-1; Site: 1: EcoRV;  
Site 2: NotI; RNA obtained from three placentas from  
female C57/BL6 mouse at 16 days pregnancy. Tissues were  
snap-frozen and kept at -80C for two days before RNA  
extraction and purification (Tri-reagent method). cDNA was  
primed using oligo-dT primer:  
5'-GCACGTAGTCTGATCGGAGCGCGGCC(T)25-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >1kb  
resulted in an average insert size of 1.3 kb. This  
primary, microquantity library is normalized to Cots  
(non-normalized primary library is NIH MGC 222) and was  
constructed by Express Genomics (Frederick, MD)"

## ORIGIN

Alignment Scores:	
Pred. No.:	1,828-94
Score:	880.50
Percent Similarity:	87.32%
Best Local Similarity:	85.85%
Query Match:	79.68%
DB:	14
Length:	773
Matches:	176
Conservative:	3
Mismatches:	11
Indels:	16
Gaps:	1

US-09-817-198C-2 (1-212) x CK030691 (1-773)

Qy	8	LeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeuLeuCysArg	27
		:	
Db	86	CTGTTTCTCTGCTTAAAGACTTAACTGGAATGGATCTGACTGCCTAAA-GCA---	141
Qy	28	PheThrAspAsnGluPheHisSerSerHisIleSerThrIleGlyValAspPheLysMet	47

```

142 Db -----|
48 Qy LysThrIleGluValAspGlyIleLysValArgIleGlnIleTyrAspThrAlaGlyCln 67
160 Db AAGACTATCGAAGTAGACGCGCATCAAGTGAATACAGATTGGGACACAGCGGCAG 219
68 Qy GluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuVal 87
220 Db GAGAGGTACACAGACTATCACAAAGCAGTACTATCGCGGAGCCAGGGAATATTTTAGTC 279
88 Qy TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyrValSerAspValAsp 107
280 Db TACGACATTAGCAGTGAAGCGCTCTCATCAGCATATCATGAAGTGGTCAAGTGGAT 339
108 Qy GluTyrAlaProGluGluValGlnLysIleLeuIleGlyAsnLysAlaAspGluGluCln 127
340 Db GAGTACGCTCCAGAAGAGTCCAGNAGATCTTAATGGGATATAGGCTGATGAAGAGCAG 399
128 Qy LysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyrGlyMetAspPhe 147
400 Db AAACGGCAGGTGGGAGAGAGCAGGGGCGAGCGCTGGCTAAGGAGTACCGCATGCACTTC 459
148 Qy TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167
460 Db TACGAACAAGTGGCTGCACCAACCTTAATATTAAGAGTCTCTTCACTCGTCTGACGGAG 519
168 Qy LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetAtqAlaSerAsnGlu 187
520 Db CTGGTGTCTGCAGGCCACAGAAAGAGCTGGATGTCTCCGAAACACAGTGCACGACAG 579
188 Qy LeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSer 207
580 Db CTGCACCTGGCGAGCTGGAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639
208 Qy LysThrCysTyrCys 212
640 Db AAGACCTGCTGGTGC 654

RESULT 12
AY420961 515 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM7398 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420961
VERSION AY420961.1 GI:39776918
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
location/Qualifiers
1..515
/organism="Pan troglodytes"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>515
/locus_tag="HCM7398"

gene
ORIGIN
Alignment Scores:
Pred. No.: 3.31e-94 Length: 515
Score: 876.00 Matches: 169
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 79.28% Indels: 0
DB: 29 Gaps: 0

US-09-817-198c-2 (1-212) x AY420961 (1-515)
Qy 43 ValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTyr 62
Db 3 GTTACATTTTAAAGTAGAAGACCATAGAGTACCGGCATCAAGTGGGATACAGATCTGG 62
Qy 63 AspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGln 82
Db 63 GACACTCGAGGCGAGGAGATACACAGCATCACAAAGCAGTACTATCGCGGCGCCAG 122
Qy 83 GlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyr 102
Db 123 GGGATATTTTGGTNNNCGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGG 182
Qy 103 ValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLys 122
Db 183 GTCACTGACCTGGATGATAGTACGCCACCAAGAGCGCTCCAGAAAGATCTTATTGGGAATAG 242
Qy 123 AlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnGlnLeuAlaLysGlu 142
Db 243 GCTGATGAGGAGCAGAAACCGGAGTGGGAGAGAGAGAGGAGGAGGAGGAGGAGGAG 302
Qy 143 TyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPhe 162
Db 303 TATGGCATGGACTTCTATGAACAAAGTGCCTGCACCAACCTCAACATTAAGAGTCAATC 362
Qy 163 ThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeuArgMet 182
Db 363 ACGCGTCTGACAGAGCTGGTGTCTGACGGCCCATAGGAGGAGCTGGAGGCGCTCCGGATG 422
Qy 183 ArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGly 202
Db 423 CGTGCCAGCATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGG 482
Qy 203 ProAlaAsnSerSerLysThrCysTyrCys 212
Db 483 CCAGCGAACTCTTCGAAACCTGCTGCTGTC 512

RESULT 13
AY420962 515 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM7398 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420962
VERSION AY420962.1 GI:39776919
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302

```

2 (bases 1 to 515)  
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence as made by sequencing genomic exons and ordering them  
 based on alignment.

FEATURES  
 source  
 1..515  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>515  
 /locus\_tag="HCM7398"

gene  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9,96E-94 Length: 515  
 Score: 872.00 Matches: 167  
 Percent Similarity: 99.41% Conservatives: 2  
 Best Local Similarity: 98.24% Mismatches: 1  
 Query Match: 78.91% Indels: 0  
 DB: 29 Gaps: 0

US-09-817-198C-2 (1-212) x AY420962 (1-515)

Qy	43	ValAspPheLysMetIysThrIleGluValAspGlyIleLysValArgIleGlnIleTrrp	62
Db	3	GTTCACCTTTAAGATGAAGACTATCGAAGTAGACGGCATCAAGTCAGATACAGATTGG	62
Qy	63	AspThrAlaGlyGlnGluArgTyrGlnThrIleThrIysGlnTyrTrpArgAlaGln	82
Db	63	GACACAGCAGGGCAGGAGAGGTACAGACTATCACAAAGCAGTACTATCGCCGAGCCAG	122
Qy	83	GlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrrp	102
Db	123	GGATATATTTTAGTCTACGACATTAGCAGTAGCGCTCCTATCAGCATATCATGAAGTGG	182
Qy	103	ValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLys	122
Db	183	GTCACTGACGTGGATAGTAGTCGCTCCAGAAAGAGTCCAGAAGATCTTAATTGGGAATAAG	242
Qy	123	AlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAlaLysGlu	142
Db	243	GCTGATGAAGACAGAAACCGCAGGTGGGAGAGACAGCGGGCAGCAGCTGGCTTAAGGAG	302
Qy	143	TyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPhe	162
Db	303	TACGGCATGACCTTACAGAAACAGTGCCTGCACCAACCTTAATATAAGAGTCTCTTC	362
Qy	163	ThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMet	182
Db	363	ACTCGTCTGACGAGCTGGTGCTGCAGGCCACAGAGAAAGAGCTGGATGCTCTCCGAACA	422
Qy	183	ArgAlaSerAsnGlnLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGly	202
Db	423	CGTGCCACGACACAGCTCGCACTGGCCGAGCTGGAGGAGGACCAAGCAAACTCGAGGGC	482
Qy	203	ProAlaAsnSerSerLysThrCysTrrpCys	212
Db	483	CCAGCAAACTCTTCAAAAGACCTGCTGTGTC	512

RESULT 14  
 BF966292  
 LOCUS  
 DEFINITION  
 6022866292F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4375366 5',  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 BF966292.1 GI:12333507  
 KEYWORDS  
 EST.



```
Db 309 AAGTGGGTCTAGTACGTGGATGAGGTAGGAGATGCCACCTCAGTCGCCGGGTGTGGAGAG 368
Qy 108 ----- 108
Db 369 GGTGCTTCACCGGGAAGCCAGCGGAGGCCAGATGGGAGGCAATGCTTCCAGGAAG 428
Qy 109 -----TyrAlaProGluGlyValGlnLysI 117
Db 429 CTTGCTTCCACAGCCCTGGATGAAGACCTCTGGTACGCACAGAGGCGTCCAGAAAGAT 488
Qy 117 eLeuileGlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyG1 137
Db 489 CTTATTGGGAATAAGCTGATAGGAGCAGAAACCGCAGGTGGGAGAGAGAGCAAGGCCA 548
Qy 137 nGlnLeuAlaLysGluTyrGlyMetAsp-PheTyrGluThrSerAlaCysThr-AsnLeu 156
Db 549 GCAGCTGGCGNAGAGATATGCGATGGACTTCTATGAACCAAGTGCCTGCACCAACCTC 608
Qy 157 AsnIleLysGluSerPheThr-ArgLeuThrGluLeuValLeuAlaHisArgLys-- 175
Db 609 AACATTAAAGAGTCATTACCGCGCTCTGACAGAGCTGGTCTGCAGGCCCATAGGAAAGG 668
Qy 176 GluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 195
Db 669 GACCTGGAAGGCCCTCCGATGCGTGCAGCAATGAGTTGGCAGCGGCAAGCTGGAGGAAC 728
Qy 196 GluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTyr 211
Db 729 GAGGAGGGCCAAACCGAGGGCCAGG-AACTTTGGAAAACCTGGGTGG 775
```

## RESULT 15

```
CB578521 LOCUS CB578521 623 bp mRNA linear EST 03-APR-2003
DEFINITION AMGNVUC:NRH7-00035-B1-A nrhy7 (10850) Rattus norvegicus cDNA clone
nryh7-00035-b1 5', mRNA sequence.
```

```
ACCESSION CB578521
VERSION CB578521.1 GI:29522562
```

## KEYWORDS

EST.

## SOURCE

Rattus norvegicus (Norway rat)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 623)

Angen EST Program.

Angen Rat EST Program

Unpublished (2003)

Contact: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00035 row: b column: 1.

## FEATURES

source

1. 623

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="nrhy7-00035-b1"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat

hypothalamus adult female Wistar rat avg. Insert size 2.3

kb fraction 6 and 7"

## ORIGIN

Alignment Scores: 1.73e-90 Length: 623

Pred. NO.: 846.00 Matches: 162

Score: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 76.56% Indels: 0

DB: 14 Gaps: 0

US-09-817-198C-2 (1-212) x CB578521 (1-623)

```
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuileGlyAspSerGlyValGly 20
Db 136 ATGCGGAAACAGTACGATGTGCTATTTCGGCTACTGCTGATCGGGGACTCCGGGGTGGGC 195
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 196 AAGACCTGCGCTGCTATGTCGCGCTTCCCGACACAGAGTTCCACTCCTCGCATATCTCCACC 255
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 256 ATCGTGTGTTGACCTTTAAGATGAAACCATCGAGTAGACGGCATCAAAGTGAGGATACAG 315
Qy 61 IleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 316 ATTTGGGACACAGCAGCGGAGGAGGTACCAGACTATCACAAAACAGTACTACCGGCGA 375
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 376 GCCCAGGGAATATTTTTAGTCTACGACATTAGCAGCGAGCGCTCTTTATCAGCACATCATG 435
Qy 101 LysTyrValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 436 AAGTGGGTGAGTGCAGTGGATGATAGTACGCCCCAGAGGCGTCCAGAAAGATCTCTCATCGG 495
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db 496 AATAAGGCTCAGCAAGAAACAGAAACGCGAGGTGGGAGAGAGCAAGGCGACAGCTGGCT 555
Qy 141 LysGluTyrGlyMetAspPheThrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 556 AAGGAGTATGGCATGGACTTTTACGAAACAAGTGCCTGCACCAACCTTAACATTAAAGAG 615
Qy 161 SerPhe 162
Db 616 TCCTTC 621
```

Search completed: March 21, 2004, 13:39:16

Job time : 2507 secs

**THIS PAGE BLANK (USPTO)**